

Child Development

VOLUME 30

MARCH, 1959 to DECEMBER, 1959

Editor

WILLIAM E. MARTIN
Purdue University

Consulting Editors

ORVILLE G. BRIM, JR.
Russell Sage Foundation

JACOB L. GEWIRTZ
National Institute of Mental Health

IRWIN M. MARCUS
Tulane University Medical School

ALBERTA ENGVALL SIEGEL
Pennsylvania State University

IRVING E. SIGEL
Merrill-Palmer School

CHARLES D. SMOCK
Purdue University

CHARLES C. SPIKER
State University of Iowa

GENEVIEVE STEARNS
State University of Iowa

HAROLD W. STEVENSON
University of Texas

JAMES M. TANNER
University of London

Managing Editor

KATE HOFFMAN

CHILD DEVELOPMENT PUBLICATIONS

*Society for Research in Child Development, Inc.
Purdue University, Lafayette, Indiana*

1959

ADA
ADA
AK
AN
BAY
BEE
BEE
BE
BIJ
BL
CA
CL
DA
DU
DU
DU
EM
FI
FI
FI
G
G
G
G
H
H
H
H

CONTENTS

VOLUME 30, 1959

ADAMS, J. STACY. Advice Seeking of Mothers as a Function of Need for Cognition ..	171
ADAMSON, JEAN. (see Siegel, Alberta Engvall)	533
AKUTAGAWA, DONALD, & BENOIT, E. PAUL. The Effect of Age and Relative Brightness on Associative Learning in Children	229
ANTONOVSKY, HELEN FAIGIN. A Contribution to Research in the Area of the Mother-Child Relationship	37
BAYLEY, NANCY. (see Rheingold, Harriet L.)	363
BECKER, WESLEY C. (see Peterson, Donald R.)	119
BEILIN, HARRY. Teachers' and Clinicians' Attitudes Toward the Behavior Problems of Children: A Reappraisal	9
BENOIT, E. PAUL. (see Akutagawa, Donald)	229
BIJOU, SIDNEY W., & STURGES, PERSIS T. Positive Reinforcers for Experimental Studies with Children—Consumables and Manipulatables	151
BLUM, ABRAHAM. The Relationship Between Rigidity-Flexibility in Children and Their Parents	297
CASTANEDA, ALFRED. (see Lipsitt, Lewis P.)	273
CLIFFORD, EDWARD. Ordering of Phenomenon in a Paired Comparisons Procedure ..	381
DAVIDSON, KENNETH S. Interviews of Parents of High Anxious and Low Anxious Children	341
DURKIN, DOLORES. Children's Acceptance of Reciprocity as a Justice-Principle	289
DURKIN, DOLORES. Children's Concepts of Justice: A Comparison with the Piaget Data	59
DURRETT, MARY ELLEN. The Relationship of Early Infant Regulation and Later Behavior in Play Interviews	211
EMMERICH, WALTER. Young Children's Discriminations of Parent and Child Roles ..	403
FISHER, RHODA LEE. (see Fisher, Seymour)	389
FISHER, SEYMOUR, & FISHER, RHODA LEE. A Developmental Analysis of Some Body Image and Body Reactivity Dimensions	389
FREEDHEIM, DONALD K., & REICHENBERG-HACKETT, WALLY. An Experimental Investigation of Parent-Child Attitudes with the PARI Scales	353
GOVATOS, LOUIS A. Relationships and Age Differences in Growth Measures and Motor Skills	333
GRATCH, GERALD. The Development of the Expectation of the Nonindependence of Random Events in Children	217
GRAY, SUSAN W. Perceived Similarity to Parents and Adjustment	91
GRIMES, ETHEL T. (see Whitacre, Jessie)	177
HAFNER, A. JACK, & KAPLAN, ARTHUR M. Children's Manifest Anxiety and Intelligence	269
HAFNER, A. JACK, & KAPLAN, ARTHUR M. An Item Analysis of the Children's Manifest Anxiety Scale	481
HARRIS, DALE B. Florence L. Goodenough, 1886-1959	305
HARRIS, DALE B. Sex Differences in the Life Problems and Interests of Adolescents, 1935 and 1957	453

344815

CHILD DEVELOPMENT

HARTUP, WILLARD W. An Evaluation of the Highberger Early-Adjustment-to-School Scale	421
HELLMER, LEO A. (see Peterson, Donald R.)	119
HITCHCOCK, ETHEL ALICE. (see Siegel, Alberta Engvall)	533
HOLTON, RUTH B. (see Spiker, Charles C.)	471
KAGAN, JEROME, & MOSS, HOWARD A. Parental Correlates of Child's IQ and Height: A Cross-Validation of the Berkeley Growth Study Results	325
KAPLAN, ARTHUR M. (see Hafner, A. Jack)	269, 481
KEMBLE, JOAN DANSKIN. (see Lipsitt, Lewis P.)	273
KING, F. J. (see Phillips, Beeman N.)	253
KOHN, LYNETTE GAYLE. (see Siegel, Alberta Engvall)	131
LARK-HOROVITZ, BETTY, & NORTON, JAMES. Children's Art Abilities: Developmental Trends of Art Characteristics	433
LEVY, EDWIN. Children's Behavior Under Stress and Its Relation to Training by Parents to Respond to Stress Situations	307
LEVY, NISSIM. (see Lipsitt, Lewis P.)	547
LIEBERT, ROBERT S., & RUDEL, RITA G. Auditory Localization and Adaptation to Body Tilt: A Developmental Study	81
LIPSITT, LEWIS P., CASTANEDA, ALFRED, & KEMBLE, JOAN DANSKIN. Effects of Delayed Reward Pretraining on Discrimination Learning of Children	273
LIPSITT, LEWIS P., & LEVY, NISSIM. Electrotactual Threshold in the Neonate	547
MACCOBY, ELEANOR E. Role-taking in Childhood and Its Consequences for Social Learning	239
MCGUIRE, CARSON. (see Phillips, Beeman N.)	253
MOSS, HOWARD A. (see Kagan, Jerome)	325
NAKAMURA, CHARLES Y. The Relationship Between Children's Expressions of Hostility and Methods of Discipline Exercised by Dominant Overprotective Parents	109
NORTON, JAMES. (see Lark-Horovitz, Betty)	433
OLTEAN, MARY. (see Zuckerman, Marvin)	27
OWEN, WAYNE A. Effects of Motivating Instructions on Reaction Time in Grade School Children	261
PALERMO, DAVID S. Racial Comparisons and Additional Normative Data on the Children's Manifest Anxiety Scale	53
PETERSON, DONALD R., BECKER, WESLEY C., HELLMER, LEO A., SHOEMAKER, DONALD J., & QUAY, HERBERT C. Parental Attitudes and Child Adjustment	119
PHILLIPS, BEEMAN N., KING, F. J., & MCGUIRE, CARSON. Studies on Anxiety: I. Anxiety and Performance on Psychometric Tests Varying in Complexity	253
POFFENBERGER, THOMAS. A Research Note on Father-Child Relations and Father Viewed as a Negative Figure	489
QUAY, HERBERT C. (see Peterson, Donald R.)	119
REICHENBERG-HACKETT, WALLY. (see Freedheim, Donald K.)	353
RHEINGOLD, HARRIET L., & BAYLEY, NANCY. The Later Effects of an Experimental Modification of Mothering	363
ROSENBERG, B. G., & SUTTON-SMITH, B. The Measurement of Masculinity and Femininity in Children	373
ROSENBLITH, JUDY F. Learning by Imitation in Kindergarten Children	69
RUDEL, RITA G. (see Liebert, Robert S.)	81
SCREVEN, C. G. Research on Running Time and Physical Work of Children Under Various Reinforcement Conditions	461
SHOEMAKER, DONALD J. (see Peterson, Donald R.)	119
SIEGEL, ALBERTA ENGVALL, & KOHN, LYNETTE GAYLE. Permissiveness, Permission, and Aggression: The Effect of Adult Presence or Absence on Aggression in Children's Play	131

INDEX

SIEGEL, ALBERTA ENGVALL, STOLZ, LOIS MEEK, HITCHCOCK, ETHEL ALICE, & ADAMSON, JEAN. Dependence and Independence in the Children of Working Mothers	533
SIMON, MARIA D. Body Configuration and School Readiness	493
SMILLIE, DAVID. An Evaluation of the Channel System on the Baby Grid	279
SPIKER, CHARLES C. Performance on a Difficult Discrimination Following Pretraining with Distinctive Stimuli	513
SPIKER, CHARLES C., & HOLTON, RUTH B. Similarity of Stimuli and of Responses in the Successive Discrimination Problem	471
SPIKER, CHARLES C., & WHITE, SHELDON H. Differential Conditioning by Children as a Function of Effort Required in the Task	I
STEVENSON, HAROLD W. (<i>see</i> Weir, Morton W.)	143
STOLZ, LOIS MEEK. (<i>see</i> Siegel, Alberta Engvall)	533
STURGES, PERSIS T. (<i>see</i> Bijou, Sidney W.)	151
SUTTON-SMITH, B. (<i>see</i> Rosenberg, B. G.)	373
WEIR, MORTON W., & STEVENSON, HAROLD W. The Effect of Verbalization in Children's Learning as a Function of Chronological Age	143
WHITACRE, JESSIE, & GRIMES, ETHEL T. Some Body Measurements of Native-Born White Children of Seven to Fourteen Years in Different Climatic Regions of Texas	177
WHITE, SHELDON H. (<i>see</i> Spiker, Charles C.)	I
ZUCKERMAN, MARVIN. Reversed Scales to Control Acquiescence Response Set in the Parental Attitude Research Instrument	523
ZUCKERMAN, MARVIN, & OLTEAN, MARY. Some Relationships Between Maternal Attitude Factors and Authoritarianism, Personality Needs, Psychopathology, and Self-Acceptance	27

1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1940
1941
1942
1943
1944
1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025

DIFFERENTIAL CONDITIONING BY CHILDREN AS A FUNCTION OF EFFORT REQUIRED IN THE TASK

CHARLES C. SPIKER* and SHELDON H. WHITE¹

State University of Iowa

Usually, discrimination problems are scored by counting correct responses or errors, and the effects of experimental variations on discriminative behavior, by comparing these measures. However, for purposes of analysis at least, we recognize that behavior in a discrimination problem involves two response tendencies, those to the positive and negative stimuli. When a subject increases his percentage of correct choices, this could be due to an increased tendency to respond to the positive stimulus and/or a decrease in the tendency towards the negative stimulus.

In differential conditioning, *S* does not make a choice. Two stimuli are alternately presented, and some measurable aspect of a uniform response is taken as an index of the strength of the tendency to respond to each. Thus, if one is reinforced and the other is not, we may obtain independent, trial-by-trial measures of changes in response tendencies to a positive and to a negative stimulus. The plausible assumption that these measures are related to the choice behavior of the simultaneous discrimination problem has received some support in animal research, in special procedures where these two paradigms are mixed (4).

The present study deals with the effects of effort on children's performance on a differential conditioning problem. For one group (Heavy, *H*), responding required a force of two-thirds of a maximum effort previously

* Iowa Child Welfare Research Station, State University of Iowa, Iowa City.

¹ The second author is now at the University of Chicago.

CHILD DEVELOPMENT

obtained from S; for the other (Light, L), it required one-third of that effort. With the heavier load, considerable effort would be involved in responding and, presumably, a higher premium would be placed on successful differentiation between stimuli which are or are not accompanied by reinforcement. One might then expect Group H to give a smaller percentage of its responses to a nonrewarded stimulus than would Group L.

METHOD

Subjects

Children from the Iowa Child Welfare Research Station Preschools were used. Some had had experience with this or other research apparatus, while others were naive. Experienced Ss averaged 57 months of age, and naive Ss 56 months. Half of each experience group was assigned, by lot, to Group L and the other half to Group H. Six children were not used because they failed to respond regularly (more than once or twice per trial) during pretraining. Because one S in Group L was not available for the second day of the procedure, the data of a randomly chosen S in Group H was not used. The remaining N of 42 was composed of 21 Ss, 13 experienced and 8 naive, under each effort condition.

Apparatus

The apparatus has been described in detail elsewhere (5). It allowed free responding within discrete trials. At eye level for the child was a circular opal-flash window 3 in. in diameter, through which the stimuli, a deeper or a lighter hue of red, were presented. To the lower right was a handle, to the lower left a pint plastic cup in which marbles were collected. Just above the cup, a white 7-W bulb flashed briefly each time a marble was delivered. The handle was attached to a microswitch in the control circuit through a spring, which could be adjusted to require a more or less forceful pull. As a measure of the force required, a scale was used to draw out the handle and the scale's reading was taken when the click of the microswitch was heard.

A stepping switch presented the sequence of positive and negative stimuli. Two Hunter timers provided a uniform stimulus duration (3 sec.) and interstimulus interval ($2\frac{1}{2}$ sec.). A switch enabled E to choose between marble delivery to both lights or reinforcement of the positive light only.

Procedure

Before the first experimental session, the handle of the scale was offered to each S at chest height, and he was encouraged to pull as hard as he could. The readings for three tries were combined to give an average weight pull (AWP) score for each S. Then, before he was brought to the experimental room on each training day, the spring on the response handle was set to require the predetermined one-third or two-thirds of his AWP.

On each day, *S* first selected a toy which was to be won in the "game," and was instructed that he was to obtain as many marbles as he could in order to win it. On his first day, he was familiarized with the apparatus. The two stimuli were alternated irregularly. The *E* pulled the handle on two trials to demonstrate the apparatus, and *S* was encouraged to pull repetitively and rapidly on six trials, receiving a marble for each pull. The *E* then went to the control room and 48 differential conditioning trials were given, during which only responses to the deeper red stimulus were reinforced. Each consecutive block of six trials was made up of three positive and three negative stimulus presentations, randomly assorted. The number of handle pulls on each trial was recorded.

One day, or occasionally two or three, intervened before a second experimental session. The *E* then instructed *S* that this was the same game and 48 additional differential conditioning trials were given.

RESULTS

Percentage of Rewarded Responses

A "differentiation" measure, roughly analogous to the percentage-choice measure in the simultaneous discrimination problem, was computed in the following manner: for each treatment subgroup (each combination of one

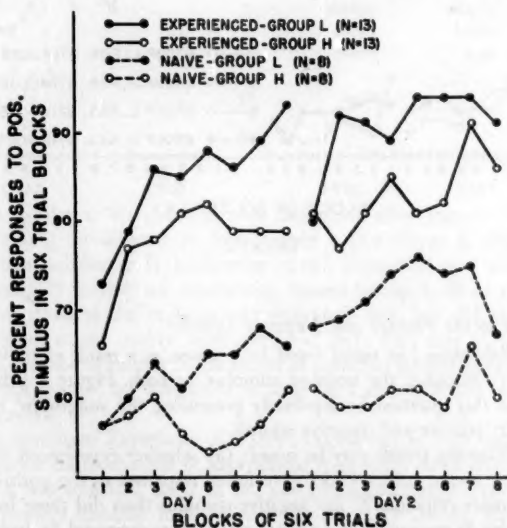


FIGURE 1

CHILD DEVELOPMENT

experience level with one effort condition), the percentage of responses in a six-trial block which were given during the three trials to the positive hue was computed. If the Ss of that subgroup had favored neither stimulus in their responding—that is, had not differentiated between the stimuli—the percentage would equal 50.

From Figure 1 it is apparent that: (a) over trials, all subgroups showed improved differentiation in favor of the positive stimulus; (b) Ss in Group L, allowed to respond with lesser effort, differentiated better than Ss in Group H; (c) experienced Ss differentiated better than did naive Ss.

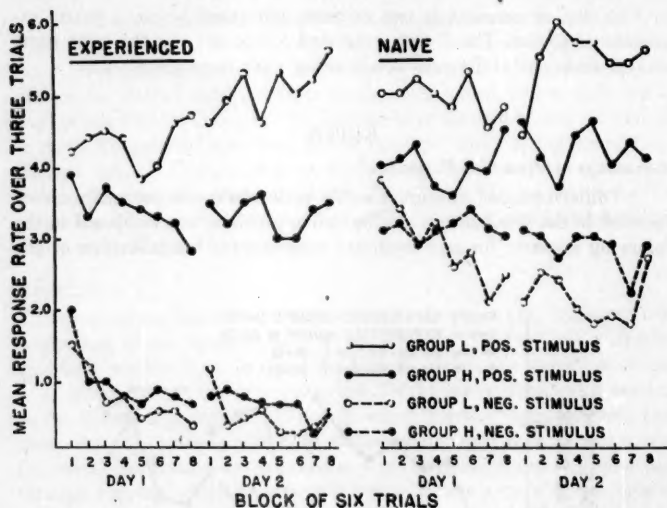


FIGURE 2

Responses to the Positive and Negative Stimuli

The differences just noted could have arisen as a result of responses to the positive stimulus, the negative stimulus, or both. Figure 2 aids in the analysis of this question by separately presenting the subgroups' response rates to the positive and negative stimuli.

The following trends may be noted: (a) whether experienced or naive, Ss working under greater effort gave fewer responses to the positive stimulus and more responses to the negative stimulus than did those in Group L; (b) naive Ss responded at a greater rate than experienced Ss, irrespective of whether heavier or lighter effort, or the positive or negative stimulus, is

CHARLES C. SPIKER and SHELDON H. WHITE

being considered; (c) while naive Ss gave only slightly more responses to the positive stimulus than did experienced Ss, they gave a great many more responses to the negative stimulus.

Statistical confirmation of these observed trends was found in an analysis of variance (2), which is summarized in Table 1. The significant Experience effect reflects the finding that naive Ss responded at a greater rate than did experienced Ss. The significant Stimuli effect indicates that, overall, the Ss in this experiment differentiated reliably in favor of the positive stimulus.

TABLE 1

SUMMARY OF ANALYSIS OF VARIANCE OF TOTAL RESPONSES TO THE CORRECT AND INCORRECT STIMULI

Source	df	Mean Square	F	p
Subjects	41			
Weight (W)	1	10519	1.46	.20
Experience (E)	1	79949	11.11	.005
W × E	1	777
error (b)	38	7199		
Within Subjects	42			
Stimuli (S)	1	366696	142.79	.001
W × S	1	27723	10.80	.005
E × S	1	21216	8.26	.01
W × E × S	1	0
error (w)	38	2568		
Total	83			

The significant Weight × Stimuli interaction effect may be interpreted as confirming the observation, from Figure 1, that Group L differentiated better than did Group H. Evaluation of this interaction by *t* tests showed that Group H Ss did not significantly exceed Group L Ss in number of responses given to the negative stimulus ($p > .20$), but did give significantly fewer to the positive stimulus ($p < .05$). This may be interpreted to mean that the difference in the extent of the effort groups' differentiations is primarily caused by differences in their rates of responding to the positive stimulus.

The significant Experience × Stimuli interaction, in turn, suggests that the superior differentiation of experienced over naive Ss is reliable. Using *t* tests, it was seen that naive Ss did not give significantly more responses to the positive stimulus than did experienced Ss ($p > .20$), but did give more to the negative stimulus ($p < .001$). The difference between these two conditions seems to lie in rate of responding to the negative stimulus.

CHILD DEVELOPMENT

DISCUSSION

The superior differentiation of Group L compared with Group H is in opposition to the tentative expectation with which the experiment was begun.

The result was largely produced by the lower rate of response to the positive stimulus emitted by the Group H Ss. Either the two effort groups differed in their tendency to respond to the positive stimulus or else their basic response tendencies were the same and an artificial ceiling was placed on Group H's responding by the greater skill and effort required to work their response handle. The choice is, thus, analogous to one between habit versus performance factors, and might best be made by transfer experiments of conventional design, in which handle weights are shifted up or down in the course of training.

Although the observed trend is not statistically reliable, it is interesting that Group H gave slightly more, rather than fewer, responses to the negative stimulus than did Group L. The added weight, if anything, caused Ss in Group H to do more work, in apparent violation of the principle of least effort. If this result can be repeated, then it would appear necessary to investigate the role of skill learning in a task such as the present one—that is, the extent to which Ss can learn to make fairly precise and effortful patterns of responses while at the same time using those responses as instruments in another learning.

The percentage measure indicates that experienced Ss differentiated between the stimuli better than did naive Ss. If response tendencies to the positive and negative stimuli are examined separately, however, it is seen that the naive Ss tended to give somewhat more responses to the positive stimulus. This would imply that the differences owing to experience are not principally due to deficiencies in the naive Ss' ability to manipulate the response handle although, possibly, their higher rate was produced at the expense of greater concentration on the mechanics of pulling.

The chief difference between the two experience groups is found in their response to the negative stimulus. Experienced Ss were considerably more successful than naive Ss in reducing their response rate to the nonreinforced stimulus. If we make the reasonable assumption that the forming of a differentiation is a process quite similar to that of forming a discrimination, then this difference may be attributed to nonspecific transfer from previous discriminations learned at the preschools by experienced Ss. A discrimination learning set has been produced in preschool children (1, 3)—that is, it has been shown that the learning of one discrimination facilitated the learning of a later one in a manner which could not be accounted for by the transfer of specific or generalized habits. The skills which are transferred have not been established, but might conceivably include orienting, labelling, or rehearsal practices.

In conclusion, it might be noted that the findings suggest that two conditions producing superior differentiation, lesser effort and more experi-

ence, do so in a demonstrably different manner—one by increasing responses to the reinforced stimulus, the other by decreasing responses to the non-reinforced stimulus. Some empirical support is thus offered for the usefulness of analyzing children's choice learning in terms of separate response tendencies even in situations where these tendencies cannot be separately observed.

SUMMARY

This experiment studied the effects of effort and experience upon preschool children's learning of a differential conditioning problem. In a procedure where Ss responded freely during discrete trials, two stimuli were alternated in irregular sequence, with responding to only one of them reinforced by marble delivery. Each S in Group L had the response handle adjusted to require a force of one-third of a maximum effort previously obtained from him. Ss in Group H had the response handle adjusted to two-thirds of their maximum. Subgroups of experienced and naive Ss were distinguished. It was found that a higher percentage of the responses of Group L Ss in comparison with Group H Ss, and of experienced Ss in comparison with naive Ss, was given to the positive stimulus. These findings were analyzed using response rates to the positive and negative stimuli taken separately.

REFERENCES

1. HARLOW, H. F. The formation of learning sets. *Psychol. Rev.*, 1949, 56, 51-65.
2. LINDQUIST, E. F. *Design and analysis of experiments in psychology and education*. Boston: Houghton Mifflin, 1953.
3. SHEPARD, W. O. Learning set in preschool children. *J. comp. physiol. Psychol.*, 1957, 50, 15-17.
4. SPENCE, K. W. *Behavior theory and conditioning*. New Haven: Yale, 1956.
5. SPIKER, C. C. The effects of number of reinforcements on the strength of a generalized instrumental response. *Child Develpm.*, 1956, 27, 37-44.

TEACHERS' AND CLINICIANS' ATTITUDES TOWARD THE BEHAVIOR PROBLEMS OF CHILDREN: A REAPPRAISAL

HARRY BEILIN*¹

Brooklyn College

The contrast between contemporary American education and that of 40 or 50 years ago is striking in at least one respect, the influence of psychology, in particular, clinical psychology.² It would require little effort to detail the many and diverse ways teacher training, parent education and child care reflect the consequences of psychology's influence. The future historian will undoubtedly dwell upon the part played by E. K. Wickman's 1928 Commonwealth Fund monograph, "Children's Behavior and Teachers' Attitudes" in this development (51). Wickman's report, which contrasts teachers' and "mental hygienists'" attitudes toward the behavior problems of children evoked an assault upon the teacher's mode of dealing with children when it made evident that teachers' attitudes were widely at variance with those of clinicians. The effect of its publication is still felt. The contiguity of events might suggest to some a causal relationship between widespread knowledge of the Wickman findings and the emergence of clinical psychology as a force in contemporary education. However, a more temperate and realistic appraisal would accept the ubiquitous penetration of psychology into American life and not as a condition unique to education. The impact of the monograph was in actuality only one of a series of challenges to the values and attitudes of educational personnel. But ir-

* Department of Education, Brooklyn College, Brooklyn 10, New York.

¹ The author is indebted to Martin Hamburger and Louis Rosenzweig for their critical reading of this paper.

² Although the period following the first World War saw the impact of the testing movement and the effects of Behaviorism, it was not till the 1930's and 1940's that clinical psychology became a part of the child development and educational scene in a major way.

CHILD DEVELOPMENT

respective of its true role—whether as reflection, or as initiator of a *Zeitgeist*—the place of the Wickman study in education and child psychology has been significant and will probably continue to be.

Wickman's results, in the main, suggested that mental hygienists were primarily concerned with *withdrawing* and other nonsocial forms of behavior in children of elementary school age, whereas teachers of these same children were more concerned with *classroom management, authority, and sex problems*. The results influenced many (starting with Wickman) to urge teachers to adopt a hierarchy of attitudes closer to that of the clinician. This view presumed that the clinician's judgment should be accepted as the criterion for adequate and inadequate behavior. Few have challenged this thesis.

The intent of the present review is to examine what the result of 30 years of research suggests for continued acceptance of this point of view. To anticipate, it will be suggested that Wickman's findings be reinterpreted and his prescription for change in educational policy modified.

THE WICKMAN STUDY

The Wickman studies were begun in Minneapolis in 1924, but a more ambitious program was undertaken in Cleveland in 1925-1926, where the following was done:

1. In a single pilot school, teachers' characterizations of undesirable behavior, with indications of "sensitiveness" to their occurrence, were elicited by questionnaire.
2. Teachers' attitudes toward various types of problems were obtained by three measures (detailed in part below).
3. On rating scales, teachers noted their reactions to problems themselves, then to pupils in whom the problems were observed, and finally to the total adjustment of their pupils.
4. Subsequently, the teachers from 13 schools in six communities and two additional teacher groups enrolled in graduate school were studied using the rating scale method developed for the pilot study. The most important feature of this involved the rating of a number of behaviors obtained from the teachers' original freely-given characterizations of problem behavior. The results are reported as mean ratings and rankings of mean ratings.
5. Finally, 30 "mental hygienists" (8 psychiatrists, 4 psychologists, 13 psychiatric social workers and 5 teachers with social work background) from child guidance clinics in three cities were studied for their attitudes toward 50 of the same behaviors rated by teachers. The mean ratings and rankings of ratings were then contrasted and correlated with those of teachers.

The rating instructions for teachers stressed: (a) *present* problems, (b) "seriousness" of the problems or "difficulties" created by them, and

HARRY BEILIN

(c) rapid responses to the rating scale. With clinicians, the emphasis was on (a) relevance of the problem behavior for *future* adjustment, (b) though "seriousness" and "difficulty" were retained, the focus was on the "importance" of the behavior, and (c) no time limit was imposed for response to the rating scale.

The principal results can be summarized as follows:

1. Teachers were most aware of overt and aggressive behaviors, inattention to school tasks, and behaviors which violated their standards of morality. They were much less concerned with behaviors indicative of social or emotional maladjustment not directly related to school routine.

2. Boys were reported more frequently than girls for behavior problems.

3. Teachers preferred the less active, more compliant behavior of girls to the more aggressive behavior of boys. Desirable conduct for teachers, then, took on the distinguishing characteristics of girl behavior.

4. "Mental hygienists" considered withdrawing and other nonsocial forms of behavior most serious and discounted the teachers' stress on antisocial behavior and violations of school rules.

5. There was a rank order correlation of $-.22$ between the rankings by mental hygienists of 50 behavior problems and the original Cleveland teachers' ($N = 28$) rankings of the same behaviors. The correlation was $-.11$ when the full sample was used ($N = 511$).

These findings were interpreted by Wickman in Thorndikian stimulus-response terms. The teachers distinguish, he said, between the attacking and withdrawing types of behavior problems. Their attitudes are principally determined, however, by the attacking nature of the child's conduct. The aggressive behaviors are identified and considered more serious because the teacher is aroused to counterattack by virtue of the frustration in him. On the other hand, the responses to withdrawing forms of behavior are modified by sympathy and protective feelings.

On the basis of these findings Wickman then proposed that:

1. Teachers' attitudes should be influenced to become more like the "ideal" clinicians. (Clinicians' attitudes are considered ideal because their judgments are (ostensibly) based upon knowledge of research in child adjustment.)

2. Teacher attitudes should be changed not by exhortation but by (a) information about child behavior through seminars and other learning experiences; and (b) practice in therapy with children.

3. Teachers' functions be less concerned with intellectual learnings and more with life adjustment.

After the appearance of the 1928 monograph some serious limitations in method and conception were pointed out by G. Watson (50) in a critical note³ which are as cogent now as when first offered. The majority of efforts

³ One writer has even wondered how the study could have been so widely and uncritically accepted with these limitations. The answer probably rests in the fact that its thesis was part of a powerfully developing movement.

CHILD DEVELOPMENT

to rectify the deficiencies have concerned only some of the criticisms. The others, however, may be of as great issue as those treated.

Watson's objections were:

1. *The procedures themselves are open to criticism.*

a. The directions given teachers and clinicians were not the same. Teachers were instructed to rank behaviors for *present* seriousness; clinicians, for *future* adjustment.

b. The time given to respond to the questionnaires was not identical. Teachers were under the control of the experimenter; clinicians were allowed an extended period to respond.

c. No definitions were given for the behavioral terms to be rated leaving to each subject the interpretation of the terms, and thus further reducing comparability of the results.

The issues raised by Watson's concern with methodology are intimately related to other criticisms.

2. *The choice of mental hygienists' attitudes toward the behavior problems of children as a criterion for evaluating teachers' attitudes toward the same problems is open to question.*

The Wickman study and others that follow (though not all) accept the clinicians' judgments as a criterion either implicitly or explicitly. Watson observes that there is no reason to suppose clinicians to be "correct" and teachers not, rather than vice versa. Wickman is questioned for not even considering this possibility.

3. *There has been too ready an acceptance of a causal relationship between withdrawing behavior in childhood and maladjustment in adulthood.*

In addition to questioning whether the term "withdrawing" means the same thing to teachers and clinicians, Watson questioned whether withdrawing behavior in childhood is causally related to, or predictive of, maladjustment in adulthood. Although this is contended in more than one theoretical position, there was very little evidence for the validity of this claim in 1928, and little more is available now. In Watson's paper there is reference to a pilot study which, for all its limitations, casts some doubt on the aforementioned assumption.

THE LITERATURE

After the Watson critique there was concern for the validity of the results and replications were undertaken with one or another modification in design, ultimately making comparability difficult.⁴

The studies, in the main, fall into the following groupings according to procedures used:

⁴ We shall consider all studies found that bear upon the problems focused on by Wickman and his critics in spite of the lack of comparability. Some studies are included which antedate Wickman because the data they offer are relevant.

HARRY BEILIN

A. *Studies employing descriptions of problem behavior.*

1. Teacher nomination of children with problems, followed by description and classification of problem behaviors (3, 5, 18, 29, 48).
 - a. In addition to all or part of the above, some use is made of a rating scale of problem behaviors (8, 15, 24, 25, 37, 47, 53, 54).
2. Teacher description of problem behavior (with no reference to specific children) from which a rating scale is developed or the descriptions themselves are used (10, 11, 22, 36, 41, 45, 51).
3. Children identified and described as problems by a social or therapeutic agency (6, 33).

B. *Studies employing the Wickman rating scales.*

1. With Wickman's directions (1, 20, 35).
2. With modifications of Wickman's directions (14, 19, 26, 38, 39, 40).

Confirmation of Wickman Findings

Early studies that made use of the Wickman scales "confirmed" what Wickman had found in that the rankings made by the teachers in other communities approximated those of Wickman's teachers. Boynton and McGaw (8) obtained a correlation of .87 between Wickman's and their own teacher ratings. Dickson (13), Laycock (22), Young-Masten (53) and Yourman (54) gave similar results as did the Epstein (15) and Snyder (37) studies. Thompson (45) offered the added information that teachers were more nearly in agreement with the rankings made by parents than with the rankings of child psychologists. In Young-Masten's study (53), in addition to securing conduct reports for nominated problem children, she observed a group of 28 problem children and a control group of the same number. From the classification of behavior records for each child observed, she obtained a statistically significant difference between the groups for behavior "which from the standpoint of the teacher, was annoying, and upset the order and peace of the classroom, and interfered with the other children" (p. 180).

Five years after the appearance of the Wickman report Bain (1) assessed the attitudes of teachers enrolled in graduate work using the original directions for teachers. The resulting rankings showed higher correlations with Wickman's mental hygienists than had been true for the original teacher sample. The correlations reflected greater teacher concern with *recessive, withdrawing behaviors* than with *active offenses* and *sex problems*.⁵

Replication of the Wickman study, employing the same directions, was undertaken in 1951 by Schruppe and Gjerde (35). They concluded that

⁵ In another part of the study Bain reports that one semester of instruction had little effect in changing attitudes toward problem behaviors. Whether the differences reported in the Bain study result from the select character of the sample is not known.

CHILD DEVELOPMENT

teachers' attitudes in 1951 agreed more closely with the "ideal criterion" (clinicians' attitudes) than did 1927 teacher's attitudes (a correlation of .57 (1951) compared with -.04 (1927) using means of teachers' and clinicians' ratings). None of the traits listed as most serious by one group was listed as least serious by the other. Wickman had five such differences. Griffiths' study (16), although not a replication of Wickman's, was similar to some features of it. He found that behavior difficulties most reported by teachers relate to violations of classroom rules and work rules. However, he indicated that teachers have changed over the years as evidenced by their awareness that a child who is easy to manage is not necessarily well adjusted.

Del Solar's (11) study is a report of interviews with parents and teachers which details the joys and problems of child rearing. It was the shy and withdrawing, rather than aggressive, behaviors which concerned this group of teachers and parents. These findings contradict Wickman and most other investigators. Del Solar's findings may result however, from the nature of the questions asked in the interview (only one, concerning liabilities of students, is cited), the small size of the sample (6 teachers rating 28 children), and the atypical nature of the subjects (all teachers had advanced study and were employed in a school of above average socioeconomic status).

We would conclude from these studies (holding the question of the validity of Wickman's methodology aside for the moment) that there is considerable evidence to indicate agreement with Wickman's original findings. Furthermore, there has been an observable shift in the intervening years in the attitudes of teachers in the direction of being more like those of "mental hygienists." In spite of greater congruence, however, a sizeable difference remains between the attitudes of teachers and clinicians toward behavior problems of children.

The Methodical Issue

Modifications as a rule have aimed at introducing uniformity in the administration of the problem rating scales. This has meant either modifying the directions given clinicians or teachers so that both groups could respond under the same stimulus conditions.

In Ellis and Miller's 1936 investigation (14) the instructions used were those originally given clinicians. In this study they were administered to teachers. With this change in method ratings of teachers correlate .49 with Wickman's mental hygienists (and .65 with Wickman's teachers). Again, the change is the result of increased realization of the seriousness of *withdrawing* and *recessive* personality traits. The investigators note, however, that teachers "still consider" *violations of general standards of morality and transgressions against authority* as the most seriously rated types of behavior.

Sparks' investigation (38) made use of both the Wickman teacher and clinician directions and scales. These were administered to teachers and

HARRY BEILIN

graduate students (in education). Teachers' ratings made in terms of seriousness for future adjustment (the mental hygienist form) were different from the kinds of ratings made when the directions stressed "troublesomeness" in the classroom, although in both cases they were still different from psychologists' ratings. The correlation of ratings (for teachers) between the original teacher form and the mental hygienist form of the scale was .05.

Mitchell's study (26) was conducted in the same cities where Wickman had done his 13 years earlier. The scales and directions were both modified, the scales but slightly. The directions given teachers and mental hygienists were the same; they were asked to rate behavior traits keeping in mind the behaviors of fifth and sixth grade children they had observed. (No such grade specification was made by Wickman.) The directions given the 1940 group were "almost" identical with those given by Wickman and these we presume were the instructions originally given clinicians. The correlation between means of ratings of 1927 and 1940 mental hygienists was .80. Some mental hygienists had apparently become more "conservative" in their ratings. They no longer considered *unsocial*, *withdrawing*, and other traits as extremely grave—in fact, no traits were now so considered. Whereas, the correlation in 1927 between the means of teachers and mental hygienists' ratings was .08 (by the rank difference method), the correlation in 1940 was .70. Mitchell interprets this to mean that either the identity in directions accounts for these results or teachers and clinicians have moved closer in their judgments.

In Stouffer's 1952 investigation (39) teachers were first given the original instrument for teachers; later, the original scale for mental hygienists. A group of mental hygienists were then administered the original mental hygienists' scale. Stouffer reports a correlation of .52 between the teachers' ranking of problem behaviors and mental hygienists' ranking of the same, employing different directions (the original Wickman procedure). A correlation of .61 was obtained when the instructions (those of mental hygienists) were the same for both groups. A rank order correlation of .87 was obtained between the rankings of Stouffer's and Wickman's mental hygienists.

Stouffer concluded that while teachers' attitudes toward the behavior problems of children have changed there has been little change in the attitudes of mental hygienists.⁶ The changes in teachers' attitudes reflect again the reduced importance of problems related to *honesty*, *sex*, *truancy*, and *classroom order*, and increased importance of *withdrawing* and *recessive* personality traits.

Stouffer's 1956 study (40) is of particular interest because it recognizes that a difference may exist between teachers of different grade levels. In contrast to his 1952 study the later one deals with secondary school teachers.

⁶ Note that Mitchell (26) suggests such a change although the correlation for the same relationship differs from Stouffer's by only .07.

CHILD DEVELOPMENT

The instructions were those for clinicians. The results were contrasted with those of elementary teachers in the prior study and with Wickman's (elementary) teachers. The rank order correlation of secondary teachers' rankings and elementary teachers' rankings was .88; between the same secondary teachers' and mental hygienists' ratings, .49. As reported in the earlier study, the correlation between elementary teachers' and mental hygienists' rankings was .61. Elementary teachers are, then, in greater agreement with mental hygienists than secondary school teachers. In terms of children's behaviors, elementary teachers are more concerned with withdrawing tendencies; secondary school teachers, with classroom management and problems related to class work and school routines.

The findings that the criteria of adjustment and maladjustment differ depending upon age and grade level is given support in a study by Beilin (3). In this instance, the procedure involves teacher nomination of maladjusted children, descriptions of their distinguishing characteristics, and content analysis of the descriptions. In general, an age trend (from elementary grades to young adulthood) was found with a concern (in elementary grades) for social-interpersonal aspects of adjustment (e.g., *withdrawal, aggressiveness, emotional instability*) to later concern (in high school) with character traits (e.g., *reliability, dependability*) and finally (young adulthood) with *achievement and integration into the community*.

In the 1957 study by Hunter (19), elementary and secondary school teachers were sent the Wickman mental hygienists' rating scale and instructions. His results are similar in direction to Mitchell's and Stouffer's though not in size of correlations. The correlation between teachers' and mental hygienists' rankings of mean ratings was .22. *Aggressiveness* is still rated more highly by teachers than by mental hygienists.

Using a procedure which raises some doubts, Ullmann (47) had teachers nominate well and poorly adjusted children who were then rated on a 144-item rating scale. Discrimination indices were computed for each item. The check list was then submitted to a group of mental hygienists who rated the extent to which each item was indicative of good or poor adjustment. The discrimination indices were correlated with the means of clinicians' ratings of the items. The correlation was .86. For favorable items it was .69; unfavorable, .50. These results are interpreted by Ullmann as confirming the Mitchell findings that teachers have moved closer to clinicians in their judgments (in spite of the considerable difference in procedure from Mitchell). Ullmann reports, however, that these correlations may be too high by virtue of some of his procedures.

At this point it appears that the differences in directions that were a part of the Wickman procedure quite clearly contributed to the differences demonstrated between teachers and mental hygienists. When this is controlled, however, differences still emerge and these are of the kind originally observed.

It is also apparent that there has been a change in the direction of greater congruence between the attitudes of teachers and clinicians. That

HARRY BEILIN

that congruence is not consistent for all levels of teachers has been made explicit in recent studies. It is likely that differences between elementary and secondary teachers have always existed vis-à-vis the matters here reviewed, but, where teacher and clinician attitudes appear to be the same, differences in meaning may still attach to the behaviors.

Teacher "Expertness"

The specific criteria employed by teachers and clinicians in assessing maladjustment in children have been mentioned. As already indicated, most investigators have shown teachers to be most concerned with children's behaviors that are *aggressive, disruptive of school routines*, or generally reflecting *lack of interest in school activities*. In addition, teachers are, or have been, less concerned with *withdrawing* and other nonsocial behaviors. Some investigators have characterized this as indicative of a middle class value pattern; e.g., *stealing* is the teacher's consistent concern in MacClenathan's (24) study. The emphasis upon these school disrupting traits has not been unanimous, however. Peck (29) found *undesirable personality traits* to be the greatest concern of the teacher, *regressive* traits somewhat less so, and *aggressive* behaviors least. Of only moderate import were *violations of school work demands*. Clark (10) differs from the usual view, too, in concluding that teachers are actually more annoyed by children's behaviors which annoy other children than by behaviors which affect teachers themselves.

In most of the cited studies it is implied or explicitly stated that the teacher is "wrong" in reacting as she does to the problems of children. Teachers have been criticized as untutored in the scientific facts concerning child development and are thus seen as generally being incapable of assessing children's adjustment.

Stewart (42) rejects this thesis. With 184 boys and 193 girls as subjects, a comparison was made between ratings of problem students and non-problem students. Identification was also attempted of those with and without "whole life" problems. From ratings of these youngsters she concludes that teachers are capable of distinguishing between problems as school problems or "whole life" problems. She insists that teachers possess much more insight into children's behavior than they are credited with by some investigators.

In spite of the few studies that report different patterns of teachers' attitudes, the hierarchy of attitudes seems to be quite close to Wickman's formulation. The Stewart report is important not so much because it rejects this hierarchy but rather in its highlighting the difference between clinicians and teachers as not being a matter of ignorance. What the difference is attributable to remains to be discussed.

Sex Differences

There seems to be universal agreement that boys are more likely to be identified as maladjusted or behavior problems than girls (6, 8, 15, 16, 18,

CHILD DEVELOPMENT

20, 25, 27, 29, 32, 37, 51, 53, 55). The proportion of boys (in contrast to girls) so identified ranges in these reports from 66 to 88 per cent. Not only is there a difference in proportion but behaviors which form the bases for these identifications are in part different for each sex (5, 15, 51). Ullmann's (47) interpretation is of some interest. It is *not*, he says, that "desirable conduct for teachers takes the distinguishing characteristics of girl behavior as suggested by Wickman but rather teachers assign girls more favorable ratings because they lack awareness of the manner in which girls are making their adjustment" (p. 39). Ullmann explains that boys' patterns of adjustment are more manifest to the observer, whereas girls deal with problems on an intrapsychic level. This interpretation is in the tradition of imputing lack of insight to teachers. Stewart's (42) results are again cogent. Her data do not suggest that teachers lack insight into their adjustment, but rather that they distinguish a different *kind* of adjustment for girls. Another study with young adults suggests the same (5). In this instance, sex differences in degree of adjustment are supplemented by differences in the types of behavior identified with the maladjustments of each sex.

Why should the nature of adjustment be different for boys and girls? Whatever the *ultimate* reasons (whether biological or social), the temptation is to say that the differences, in an *immediate* sense at least, result from different *expectations*. It is evident from the cited studies that boys and girls are expected to act in prescribed ways in our culture. The reasons girls are considered better adjusted by teachers is that teachers have certain expectations of what good adjustment *in school* should be and the prescription for girls' adjustment is more consistent with these expectations than the prescription for boys' good adjustment. As Wickman makes evident, the teacher is concerned with getting what she is teaching "across," and behaviors which facilitate this are more likely to be valued. The behaviors of girls are of this kind.

This approach is more acceptable to us, from the evidence, than the interpretation that teachers' attitudes are based on a lack of sophistication.

There is some evidence that men and women teachers evaluate the problems of children differently. Women are found by one investigator to rate problem behaviors as more serious than do men (14). Another study (19) reports, however, that specific problem behaviors are treated differently by each sex. Men teachers consider *sex* problems as less serious than do women; women consider *appearance* and *destruction of property* as less serious than do men. Others (4, 41) report similar findings, although in the former case it is emphasized that the similarities are greater than the differences.

Age and Grade Influences

The sixth grade appears to be modal for the nomination of children with problems, with the first and second grades offering the least. The fifth, seventh, and eighth grades also give the teacher some difficulty (18, 20, 25,

HARRY BEILIN

37). Difference in maladjustments of elementary and secondary school youngsters were reported early in the literature (18) and somewhat neglected till recently. Hildreth (18) observed that maladjusted elementary school children are more likely to be identified as *unstable*, *nervous*, or *shy*; the secondary school pupil, as *aggressive* or demonstrating *poor study habits*. Peck (29) finds the differential effect of sex of students in these identifications, however. Grade differences in problem type are reported by others as well (5, 40, 41). The differences found by Stouffer (40) have already been described. Griffiths (16) states that certain behavior difficulties (as reported by teachers and parents) increase with age; others decrease.

Socioeconomic Status

There are limited data relating socioeconomic status of the child to his identification as a problem. Levy (23) finds "socially high grade children have personality or emotional problems . . . children of lower classes have social problems" (p. 158). Yourman (54) reports a larger proportion of problem children are of lower socioeconomic status. Snyder (37) found that schools differentiated by their level of socioeconomic status yielded different numbers of problems with more from the lower groups. There was no statistically significant difference, however, in socioeconomic status between a problem group and a control group.

The study by Griffiths (16) makes the most ambitious attempt to relate socioeconomic status to the identification of behavior problems. There were few significant differences among children of different socioeconomic levels in teachers' ratings of their problems. More differences appear, however, according to the parent's ratings and the child's own ratings. Griffiths concludes that some differences exist between middle socioeconomic level children and others. In particular, they are more submissive and less aggressive.

It is apparent that few data are available as to the relationship between socioeconomic status and the behavior problems of children.

DISCUSSION

The studies reviewed suggest strongly that differences in teachers' and clinicians' attitudes existed in 1927. From that time to the present changes appear to have taken place among teachers' attitudes so that they approximate more closely those of clinicians. There is some possibility that clinicians have tempered their evaluations as well.

Despite the shift toward congruence, teachers' attitudes remain different, and different in ways not dissimilar from what they were in Wickman's day. Why? First, let us recall that Wickman and others made much of this difference. It was suggested, even insisted, that the teacher should change. Such an injunction could rest only on the premise that the clinician's attitudes were more legitimate or more correct. This view was ac-

CHILD DEVELOPMENT

cepted though G. Watson was the first and not the last to question it. Watson's position was not that the clinician was necessarily incorrect or that he should not serve as a criterion. Rather, Wickman was chided for not even considering the alternatives to accepting the clinician as criterion. For Wickman, the virtue of choosing the clinicians' attitudes as an ideal was recommended by his expert knowledge of children's adjustment. Let us examine this claim.

For one, Wickman asked clinicians to rate behaviors in light of their possible future consequences. Would the clinician, with any validity, know the future consequences of the appearance of a behavior in childhood? The answer is doubtful. In 1927 there were few if any studies which had indicated with even low degrees of certainty the outcome in adolescence or adulthood of a child's particular behavior (e.g., *withdrawing* behavior). In fact, if anything, there was some doubt that this could be done (30, 50). Evidence since then leads to even greater uncertainty (2, 21). According to some theories a withdrawn child is more likely to become maladjusted than one who is not. Yet the proof of how true this is and in what proportion for any population is almost nonexistent. Although it has been shown that in an adult *psychotic* group (52) there was a tendency for mal-adjustive behaviors to be present in childhood (information was obtained from retrospective reports), this tells us little about the prevalence of withdrawing behavior in a population of children or about the likelihood of such behaviors resulting in maladjustment, neurosis, or psychosis in adulthood.

However, there is a more important issue, in light of the functions of the therapist *qua* therapist: withdrawing behaviors present a problem to be dealt with at the time of their appearance. Such behaviors can be a basis for a visit to a therapist—in childhood as well as adulthood. The clinician is more likely to attempt some therapy than to postpone action till adolescence or adulthood (although, in some instances, this might reasonably be done). The clinician is often forced to act by immediate criteria; for he cannot wait for ultimate validation. It is thus part of his role as a therapist to be concerned about these behaviors. In essence, the behaviors with which a clinician is concerned are related to his status and the functions that accompany that status. If these behaviors were of equal relevance to the functions of the teacher, they would be equally valued. However, they are not. This has been so even in the period of "life-adjustment" programs and through the era of the "whole child," except possibly for some special groups of teachers. In spite of much pressure, teachers on the whole continue to be concerned with behaviors that facilitate or interfere with their teaching. A number of investigators recognize the difference in function between clinician and teacher (12, 18, 24, 35, 42, 50, 54) even though the teacher's role is not simple to define. It is, after all, a reflection of an educational philosophy. The prevailing philosophy of education in 1927, whether explicit or implicit, was oriented to the training of intellectual

HARRY BEILIN

skills. In the interim the function of the teacher has broadened considerably to include training in social and other skills. There has been much pressure on the teacher to be a counselor and in some ways something of a psychotherapist as well—but, at the least, to focus more on the emotional life and adjustment of the child. The question of which role is “better” is a question of values. At present, the trend is back again toward the training of intellectual skills. The trend of increasing teacher sophistication in psychology will probably continue, and will probably not revert to the level of 1927. To urge (e.g., 39, 40) that the teacher’s attitudes approximate the clinician’s is unrealistic unless the teacher’s role becomes one with the clinician—and this seems unlikely.

Other considerations recommend themselves as well. The teacher has a vital role in the socialization of the child. She is, after all, a culture carrier and to some extent a parental surrogate. Her own behaviors are significant in the child’s development of self-control, character traits, values, and work habits. These functions are certainly as important as any. There is no question that the teacher needs to be aware of withdrawing and other undesirable personality characteristics. What is questioned is the need for the teacher to concern herself with them to the same extent and in the same way as the clinician.

To summarize, the difference reported by Wickman in attitudes toward the behavior problems of children should be interpreted as reflecting differences in the roles of teachers and clinicians and the discharge of functions of the role incumbents (9, 34). The efforts of many have been directed to alter the prescription of the teacher’s role and performance in this role. This effort has in part been successful, as witnessed by the greater congruence in attitudes between teacher and clinician. In spite of the partial change in prescription, the teacher’s role remains principally task-oriented; the clinician’s, more adjustment-oriented. It seems unrealistic and possibly even undesirable to expect the teacher’s behaviors reflected in her attitudes and values to become congruent with those of clinicians. Other results reviewed here are consistent with this thesis. The reported disparity between elementary and secondary teachers results from differences in role. The high school teacher is even more subject matter-oriented than the elementary school teacher. The greatest impact of the “child-oriented” or “life-adjustment” philosophy in turn has been in the elementary school. This has resulted in a modification of role prescription for the elementary school teacher which is reflected in greater similarity between the attitudinal hierarchies of elementary teachers and clinicians. The observed differences are due not only to the teacher’s role but result from the actions of the children themselves. The pupil’s role-related behaviors change with progress through school as the youngster assumes new responsibilities and loses old ones. Behavior differences are not only bound to their age but also their sex. The same behavior is not expected or demanded of boys and girls. The attitudes of teachers in turn will reflect differences in age and sex role expectations.

CHILD DEVELOPMENT

SUMMARY AND CONCLUSIONS

The studies concerning teachers' and clinicians' attitudes toward the behavior problems of children, which have emerged principally from the initiative of the Wickman 1928 monograph, are reviewed. The following conclusions are drawn:

1. Differences existed in 1927 between the attitudes of teachers and clinicians toward the behavior problems of children. This seems to have been true in spite of the methodological limitations of the Wickman study.
2. Since 1927 there has been a shift in the hierarchy of teachers' attitudes to approximate more closely those of clinicians. This shift is not due to an artifact of research methodology. Those studies which incorporate adequate controls and consistent instructions show even greater congruence between the attitudes of the two groups.
3. There has been some change in the attitudes of clinicians although this is based upon the conclusions of one study.
4. Criteria employed in evaluating the behavior problems of children differ for elementary and secondary school teachers.
5. More boys are identified as maladjusted than girls and the criteria of maladjustment (and adjustment) differ in part for each sex.
6. The sex of the teacher affects, in part, attitudes toward children's problems.
7. Studies of the relationship of socioeconomic factors to the evaluation of children's behavior problems are inadequately dealt with in the literature.

Differences in attitudes between teachers and clinicians are interpreted in the framework of role theory. The attitudinal hierarchies of teachers and clinicians are seen as reflecting their respective roles and the ways these roles influence the organization of their respective experiences. Wickman's findings of 1927 are interpreted as indicative of the role of the teacher in that era. The role expectations of teachers have changed. Replications of the Wickman study indicate these changes have resulted in greater congruence between teachers' and clinicians' attitudes. It is suggested by virtue of the teachers' essential task-orientation and the clinicians' adjustment-orientation that complete or nearly complete congruence is not likely to be achieved.

The relationship of sex and age to attitudes toward behavior problems is also explained in the light of role theory.

REFERENCES

1. BAIN, W. E. A study of the attitudes of teachers toward behavior problems. *Child Develpm.*, 1934, 5, 19-35.
2. BEILIN, H. The prediction of adjustment over a four year interval. *J. clin. Psychol.*, 1957, 13, 270-274.

HARRY BEILIN

3. BEILIN, H. Effects of social (occupational) role and age upon the criteria of mental health. *J. soc. Psychol.*, 1958, 48, 247-256.
4. BEILIN, H., & WERNER, E. Sex differences among teachers in the use of criteria of adjustment. *J. educ. Psychol.*, 1957, 48, 426-436.
5. BEILIN, H., & WERNER, E. Sex role expectations and the criteria of social adjustment for young adults. *J. clin. Psychol.*, 1957, 13, 341-343.
6. BLANCHARD, P., & PAYNTER, R. H. The problem child. *Ment. Hyg.*, 1924, 8, 26-54.
7. BOWLES, H. H. A study of nurses' attitudes toward behavior problems in children under hospital care. *Child Developm.*, 1937, 8, 282-288.
8. BOYNTON, P. L., & MCGAW, B. H. The characteristics of problem children. *J. juv. Res.*, 1934, 18, 215-222.
9. BRIM, O. G. The parent-child relation as a social system: 1. Parent and child roles. *Child Developm.*, 1957, 28, 343-364.
10. CLARK, E. J. Teacher reactions toward objectionable pupil behavior. *Elem. Sch. J.*, 1951, 51, 446-449.
11. DEL SOLAR, C. *Parents and teachers view the child; a comparative study of parents' and teachers' appraisals of children*. New York: Bureau of Publications, Teachers College, Columbia Univer., 1949.
12. DAVIS, E. A., & MCGINNIS, E. *Parent education; a survey of the Minnesota program*. Minneapolis: Univer. of Minnesota Press, 1939.
13. DICKSON, V. E. Behavior difficulties that baffle teachers. *J. juv. Res.*, 1932, 16, 93-101.
14. ELLIS, D. B., & MILLER, L. W. Teachers' attitudes and child behavior problems. *J. educ. Psychol.*, 1936, 27, 501-511.
15. EPSTEIN, L. J. An analysis of teachers' judgments of problem children. *J. genet. Psychol.*, 1941, 59, 101-107.
16. GRIFFITHS, W. *Behavior difficulties of children as perceived and judged by parents, teachers and children themselves*. Minneapolis: Univer. of Minnesota Press, 1952.
17. HERTZMAN, J. High school mental hygiene survey. *Amer. J. Orthopsychiat.*, 1948, 18, 238-256.
18. HILDRETH, G. A survey of problem pupils. *J. educ. Res.*, 1928, 18, 1-14.
19. HUNTER, E. C. Changes in teachers' attitudes toward children's behavior over the last thirty years. *Ment. Hyg.*, 1957, 41, 3-11.
20. HURLOCK, E. B., & McDONALD, L. C. Undesirable behavior traits in junior high school students. *Child Developm.*, 1934, 5, 278-290.
21. IVES, O. L. *A critique of teachers' ratings of high school boys as an indication of later neuropsychiatric rejection for the armed services*.

CHILD DEVELOPMENT

New York: Bureau of Publications, Teachers College, Columbia Univer., 1949.

22. LAYCOCK, S. R. Teachers' reactions to maladjustments of school children. *Brit. J. educ. Psychol.*, 1934, 4, 11-29.
23. LEVY, J. A quantitative study of the relationship between intelligence and economic status as factors in the etiology of children's behavior problems. *Amer. J. Orthopsychiat.*, 1931, 1, 152-162.
24. MACCLENATHAN, R. H. Teachers and parents study children's behaviors. *J. educ. Sociol.*, 1934, 7, 325-333.
25. MCCLURE, W. E. Characteristics of problem children based on judgments of teachers. *J. juv. Res.*, 1929, 13, 124-140.
26. MITCHELL, J. C. A study of teachers' and mental hygienists' ratings of certain behavior problems of children. *J. educ. Res.*, 1942, 36, 292-307.
27. NEUMEYER, M. H. *Juvenile delinquency in modern society*. New York: D. Van Nostrand, 1949.
28. O'MALLEY, K. E. *A psychological study of the annoyances or irritations of teachers*. Unpublished doctoral dissertation, New York Univer., 1936.
29. PECK, L. Teachers' reports of the problems of unadjusted school children. *J. educ. Psychol.*, 1935, 26, 123-138.
30. PRESTON, G. H., & SHEPLER, W. McL. A study of the problems of "normal" children. *Amer. J. Orthopsychiat.*, 1931, 1, 245-256.
31. PULLIAS, E. V. How do you behave when the children misbehave? *Childh. Educ.*, 1934, 10, 230-237.
32. ROGERS, C. R. The criteria used in a study of mental health problems. *Educ. res. Bull.*, 1942, 21, 29-40.
33. ROGERS, C. R. Mental health findings in three elementary schools. *Educ. res. Bull.*, 1942, 21, 69-79.
34. SARBIN, T. R. *Role theory*. In G. Lindzey (Ed.), *Handbook of social psychology*. Cambridge: Addison-Wesley, 1954. Pp. 223-258.
35. SCHRUPP, M. H., & GJERDE, C. M. Teacher growth in attitudes toward behavior problems of children. *J. educ. Psychol.*, 1953, 44, 203-214.
36. SEIDMAN, J. M., & KNAPP, L. B. Teacher likes and dislikes of student behavior and student perceptions of these attitudes. *J. educ. Res.*, 1953, 47, 143-149.
37. SNYDER, L. M. The problem child in the Jersey City elementary schools. *J. educ. Sociol.*, 1934, 7, 343-352.
38. SPARKS, J. N. Teachers' attitudes toward the behavior problems of children. *J. educ. Psychol.*, 1952, 43, 284-291.
39. STOUTFEE, G. A. W., JR. Behavior problems of children as viewed by teachers and mental hygienists. *Ment. Hyg.*, 1952, 36, 271-285.

HARRY BEILIN

40. STOFFER, G. A. W., JR. The attitudes of secondary school teachers toward certain behavior problems of children. *Sch. Rev.*, 1956, 64, 358-362.
41. STOFFER, G. A. W., JR., & OWENS, J. Behavior problems of children as identified by today's teachers and compared with those reported by E. K. Wickman. *J. educ. Res.*, 1955, 48, 321-331.
42. STEWART, N. Teacher's concepts of "behavior problems." In *Growing points in educational research*. Washington: American Educ. Res. Ass. Rep., 1949.
43. STODDILL, R. N. Parental attitudes and mental hygiene standards. *Ment. Hyg.*, 1931, 15, 813-827.
44. STODDILL, R. N. Experiments in the measurement of attitudes toward children, 1899-1935. *Child Developm.*, 1936, 7, 31-36.
45. THOMPSON, C. E. The attitudes of various groups toward behavior problems of children. *J. abnorm. soc. Psychol.*, 1940, 35, 120-125.
46. UGER, C. The relationship of teachers' attitudes to children's problem behavior. *Sch. & Soc.*, 1938, 47, 246-248.
47. ULLMANN, C. A. *Identification of maladjusted school children*. Public Hlth Monogr. No. 7. Washington: U. S. Govt. Printing Office, 1952. (Rev. 1957)
48. WANDT, E. Measurement and analysis of teachers' attitudes. *Calif. J. educ. Res.*, 1952, 3, 10-13.
49. WANDT, E. A comparison of the attitudes of contrasting groups of teachers. *Educ. psychol. Measmt*, 1954, 14, 418-422.
50. WATSON, G. A critical note on two attitude scales. *Ment. Hyg.*, 1933, 17, 59-64.
51. WICKMAN, E. K. Children's behavior and teachers' attitudes. New York: Commonwealth Fund, 1928.
52. WITMER, H. L. Childhood personality and parent-child relationships of dementia praecox and manic depressive patients. *Smith Coll. Stud. Soc. Wk.*, 1934, 4, 287-378.
53. YOUNG-MASTEN, I. Behavior problems of elementary school children: a descriptive and comparative study. *Genet. Psychol. Monogr.*, 1938, 20, 123-181.
54. YOURMAN, J. Children identified by their teachers as problems. *J. educ. Sociol.*, 1932, 5, 334-343.
55. U. S. Children's Bureau. *Juvenile court statistics 1946-49*. Statistical Series No. 8. Washington: U. S. Govt. Printing Office, 1949.

SOME RELATIONSHIPS BETWEEN MATERNAL
ATTITUDE FACTORS AND AUTHORITARIANISM,
PERSONALITY NEEDS, PSYCHOPATHOLOGY,
AND SELF-ACCEPTANCE

MARVIN ZUCKERMAN*¹

Indiana University Medical Center

and

MARY OLTEAN¹

LaRue D. Carter Memorial Hospital

The Parental Attitude Research Inventory (PARI) consists of 23 five-item scales, each scale measuring some particular parental attitude (10). In a previous study the PARI was factor analyzed and three factors were extracted (15). These same factors emerge from factor analyses in mothers of disturbed children as well as in mothers of normals. Schaefer (9) also reports a factor analysis of the PARI using normal mothers with results substantially similar to the aforementioned study. The largest factor, Authoritarian-Control (A), seems to measure authoritarian, suppressive, punitive, and restricting attitudes. The Hostility-Rejection factor (B) can be described as measuring hostility towards children and husband, and rejection of the maternal role. The Democratic Attitude factor (C) is composed of the three positively worded scales, and seems to measure democratic attitudes. However, Democratic Attitude is independent of Authoritarian-Control and probably reflects the response tendency to agree with positive sounding generalizations more than a genuine response to the test item content. The scales composing this factor were not expected to be discriminating by the authors of the test, since they found that items stating healthy attitudes towards child-rearing in the Mark (8) and Shoben (12) studies did not differentiate effectively between the parents of normals and parents of maladjusted children. These scales were included because some parents objected to the paucity of items with which they could agree.

If the dimensions of parental attitudes measured by these scales are something more than transient opinions, changing with every new article on child-rearing in the *Ladies Home Journal*, they should be related to

* Institute of Psychiatric Research, Indiana University Medical Center, 1100 West Michigan, Indianapolis.

¹ The authors wish to thank Dr. Beatrice Barrett-Ribback and Dr. John I. Nurnberger for their careful reading of this manuscript and helpful suggestions.

CHILD DEVELOPMENT

attitudes in other areas and to relevant personality traits. Testing the relationships between personality and parental attitudes would enable one to assess the construct validity of the parental attitude test. It is possible that a parental attitude may exist, isolated from the rest of the attitudinal structure, e.g., a person may profess liberal attitudes on social issues but may demand rigid discipline in his home. However, in view of current social-personality theory and research, it is more reasonable to expect some relationship between parental attitudes and attitudes in other spheres of opinion and manifest personality traits or needs.

Authoritarianism

Adorno *et al.*, in their well-known work *The Authoritarian Personality* (1), hypothesized a relationship between social political attitudes and child-rearing behavior and attitudes. Hart (6) found a correlation of .63 between scores of mothers on the California F Scale of Authoritarianism and the mothers' self-reported use of "non-love" oriented techniques of discipline. Willis (13) found a .33 correlation between a scale measuring acceptance of authoritarian political practices and a scale measuring demands of obedience from children. Block (4) found that high-restrictive fathers, as measured by a child-rearing attitude scale with items similar to the Authoritarian-Control factor on the PARI, scored significantly higher than permissive fathers on the California E and F Scales.

Parental Attitudes and Personality

In Q-sort assessment, Block (4) found that high-restrictive fathers could be characterized as suggestible, disorganized, indecisive, effeminate, over-controlled, conforming, stereotyped, submissive to authority, and lacking confidence in themselves. In an experiment on independence of judgment Block found that a sample of this high-restrictive group showed significantly greater "yielding" behavior than a sample of his low-restrictive group.

Self-Esteem

In the Sears, Maccoby, and Levin study of parental attitudes (11), using a structured interview, two factors were found similar to factors A and B on the PARI. The Sears factor A, Permissiveness-Strictness, was similar to the PARI Authoritarian-Control factor. The Sears factor B, General Family Adjustment, was similar to the PARI Hostility-Rejection factor. A rating of the mother's self-esteem in the Sears study showed a high loading on Sears factor B and only a negligible loading on Sears factor A. Mothers with low self-esteem tended also to demonstrate low evaluation of the father, dissatisfaction with the current life situation, and low affectionate interaction with a baby.

Hypotheses

1. Authoritarian social attitudes as measured by the F scale should correlate significantly and positively with the PARI Authoritarian-Control

MARVIN ZUCKERMAN and MARY OLTEAN

factor, since this factor is supposed to measure authoritarian and suppressive child-rearing attitudes.

2. Since the PARI factor A has been conceptualized as measuring Authoritarian-Control, it was expected that this factor would relate to personality traits characteristic of the "Authoritarian Personality." One of these traits is submission to authority. We therefore would expect a positive relationship between Authoritarian Control and Deference on the Edwards Personal Preference Schedule (EPPS). This hypothesis is supported by the finding of Block (4) that his high restrictive group of fathers showed greater "yielding" behavior. The authoritarian is supposed to have attitudes of hostility and cynicism about people which would lead us to predict a negative relationship between Authoritarian-Control and Affiliation. Also characteristic of the authoritarian are repressive attitudes toward sexual impulses. Therefore, we would expect a negative correlation between Authoritarian-Control and Heterosexuality.

The PARI factor B has been described as Hostility-Rejection. The PARI scales composing this factor are Irritability (with children), Rejection of the Homemaking Role, and Marital Conflict. A woman who is high on this factor would be expected to have needs for achievement outside of the home, lack motivation to care for and devote herself to others, and show aggressiveness in interpersonal relations. Therefore a positive correlation was predicted between this factor and Achievement and Aggression and a negative correlation was predicted between the factor and Nurture on the EPPS.

3. Certain of the individual PARI scales seem to be measuring variables which are very similar to variables measured by individual EPPS scales, e.g., PARI Ascendancy of the Mother and EPPS Dominance. The difference is that the PARI attempts to measure attitudes toward home and family while the EPPS attempts to measure these attitudes in general interpersonal relations. Significant relationships between these pairs of scales would add to their construct validity.

4. The results from the Sears, Maccoby, and Levin study lead us to expect that self-acceptance would correlate negatively with the Hostility-Rejection factor.

5. Assuming that severe parental attitudes are consequences of disturbances in the personality of parents, it is expected that both parental attitude factors would correlate positively with severity of disturbance as measured by the clinical MMPI scales.

METHOD

The subject groups used were selected because other studies were being conducted with them and personality test data were available on them.

Three groups of subjects were used: (a) 60 female patients in an acute psychiatric treatment hospital, most of whom were mothers; (b) 24 mothers

CHILD DEVELOPMENT

of college students; (c) 88 student nurses, none of whom were married or mothers.

Tests used were:

1. The Parental Attitude Research Inventory (PARI), described in the introduction, was administered to all groups.

2. The California F Scale (1), a measure of authoritarian social attitudes, was given to 32 of the female patients, and to all 88 student nurses.

3. The Edwards Personal Preference Schedule (EPPS) (5), a measure of manifest personality needs drawn from the Murray system, was given to the 24 mothers of college students and 63 of the student nurses.

4. The Minnesota Multiphasic Personality Inventory (MMPI) (7), a test of psychopathological tendencies, was given to the 60 female patients.

5. A test of self-acceptance as measured by the discrepancy between self and ideal concepts was given to 63 of the student nurses. This test was devised by the senior author and consists of 21 pairs of adjectives, each pair matched on the basis of social desirability ratings. The subject must check the adjective in each pair which best describes herself. She then takes the test again checking her "ideal" description. The score of self-esteem is the number of identical responses in the ratings of self and ideal.

RESULTS AND DISCUSSION

Authoritarianism

✓ The F scale correlates significantly with the PARI Authoritarian-Control factor in the group of patients ($r = .51$) and in the student nurse group ($r = .61$) as predicted. Both correlations were significant beyond the .01 level. This would reinforce the conclusion that authoritarianism in social attitudes is an expression of a personality tendency which affects attitudes about child-rearing, as well. The possibility does exist that these correlations are due to a relationship with a common factor, the acquiescence response set (2), which may play an important role in both tests because of their similarity in structure (Likert type, unidirectional items). However, one would expect some relationship between the F scale and the other PARI factors if the acquiescence tendency accounted entirely for the relationship with factor A. The correlations between the F scale and the other PARI factors did not approach significance in either of the groups.

Personality Variables

Table 1 presents correlations between the PARI factors and selected EPPS scales. In the mothers of college students, the predictions for the Hostility-Rejection factor are borne out. The mother who tends to be hostile and rejecting in her parental attitudes tends to have a high need for achievement, a low need for nurturance, and a high need for aggression. The Authoritarian-Control factor predictions are not borne out as well. There is a negative relationship between Authoritarian-Control and the

MARVIN ZUCKERMAN and MARY OLTEAN

TABLE I
CORRELATIONS BETWEEN PARI FACTORS AND EPPS SCALES

EPPS Needs	Mothers of Students (N = 24)			Student Nurses (N = 63)		
	PARI FACTORS			PARI FACTORS		
	A	B	C	A	B	C
Achievement40	.48*	.16	.03	.20	.06
Nurturance	-.40	-.59**	.23	.00	-.10	.25*
Aggression27	.41*	-.13	.01	-.09	-.06
Affiliation	-.43*	-.67**	-.04	.09	-.10	.10
Deference23	.27	.00	.36**	.04	-.38**
Heterosexuality	-.20	-.02	.05	-.12	.19	.18

* Significant at or below the .05 level.

** Significant at or below the .01 level.

need for affiliation, but the correlation between this need and Hostility-Rejection is even higher.

In the student nurse group the correlations between parental attitude factors and personality needs are generally nonsignificant. The exceptions are positive correlations between Nurturance and Democratic Attitudes, and Deference and Authoritarian-Control, and a negative correlation between Deference and Democratic Attitudes. It is possible that parental attitudes in girls who have had no experience raising children are less a function of personality needs than in women for whom the attitudes are more than theory. One would expect that some parental attitudes are changed after there is actual experience in the raising of children. It might be expected that the changes which occur would pull parental attitudes in the direction of personality needs.

The conceptualization of factor B as a Hostility-Rejection factor is borne out by these results. A woman whose significant rewards tend to lie in achievements outside the nurturing, maternal role is one who is likely to be irritable with her children and her husband because she is functioning in a role which does not fit her needs.

The correlations between EPPS scales and PARI scales supposedly measuring similar constructs can be seen in Table 2. With the exception of the nonsignificant relationship between Suppression of Sex and Heterosexuality, these relationships are high and are all in the expected directions indicating good construct validity for these individual PARI scales.

Again the relationships between personality and parental attitude measures are not found in the student nurse group. The senior author was intending to use parental attitude measures on women pregnant with their first child in an attempt to predict the type of mother-child interaction which would ensue after birth. Perhaps the approach of studying the per-

CHILD DEVELOPMENT

TABLE 2

CORRELATIONS BETWEEN SPECIFIC PARI SCALES AND EPPS SCALES

EPPS Needs	PARI Scales	Mothers (N=24)	Student Nurses (N=63)
Achievement	13. Rejection of Home-making Role	.57**	.05
Nurturance	13. Rejection of Home-making Role	-.56**	-.08
Aggression	9. Irritability44*	-.11
Affiliation	3. Seclusion of the Mother	-.56**	.20
Dominance	19. Ascendence of the Mother51**	.11
Heterosexuality	18. Suppression of Sex	-.34	-.12

* Significant at or below the .05 level.

** Significant at or below the .01 level.

sonality traits and general attitudes of women prior to child bearing might be a more fruitful approach than trying to assess their attitudes toward child rearing.

Self-Esteem

Self-Acceptance correlated $-.37$ with Hostility-Rejection ($p < .01$). The correlations with Authoritarian-Control and Democratic Attitudes were $-.13$, and $-.15$ (not significant). The expectation derived from the Sears, Maccoby, and Levin study is borne out. This tends to reinforce the conclusions from other studies of self-acceptance (3, 14) that a person who is not accepting of self is not accepting of others, the others in this case being husband and children. Low self-acceptance can also be related to the role conflict implied in this factor. Since the mothers who score high on Hostility-Rejection are functioning in a role they do not accept, it would be difficult for them to accept themselves. One would also expect a certain amount of guilt, and consequent loss of self-acceptance, to accompany rejection of the home-making role since the maternal motive is often held up as a cultural ideal.

Psychopathology

The correlations in the patient group between the PARI factors and psychopathological tendencies as measured by the scales of the MMPI are contained in Table 3. Also included in this table are correlations between these factors and the age and education of the mother.

The age of the mother seems to have little relationship to parental attitude factors. As found in previous studies (15, 16), the education of the mother shows a marked negative relationship with the Authoritarian-Control factor. Education of the mother was closely related to her social class in the Sears, Maccoby, and Levin study (11), and both class and education were shown to influence patterns of child rearing. Emphasis on

TABLE 3

CORRELATIONS BETWEEN PARI FACTORS AND AGE, EDUCATION, AND
MMPI SCORES OF FEMALE PATIENTS (N = 60)

	A	B	C
Age12	-.04	-.21
Ed	-.51**	-.09	-.21
F27*	.44**	.01
K	-.29*	-.52**	-.05
Hs02	.25*	-.03
D	-.15	.14	-.22
Hy	-.17	-.08	-.09
Pd04	.23	-.04
Mf56*	.10	.12
Pa21	.33*	-.15
Pt17	.35**	-.06
Sc25*	.48**	-.04
Ma31*	.38**	.03
Si10	.20	-.14
Anx.09	.34**	-.08

* Significant at or below the .05 level.

** Significant at or below the .01 level.

authoritarian and suppressive attitudes in the home seems to be more characteristic of the less educated. The hostile and rejecting attitudes in factor B are unrelated to the education of the mother.

Authoritarian-Control seems to be only minimally related to areas of psychopathology as measured by the MMPI. The one outstanding exception is the pronounced relationship ($r = .56$) shown between Authoritarian-Control and the nonclinical scale, Masculinity-Femininity. When education is partialled out, this correlation drops to .44 which is still significant below the .01 level of significance. This finding is somewhat surprising since it might be expected that masculinity would bear a closer relationship with the Hostility-Rejection factor which connotes dissatisfaction with the maternal role. Perhaps the "tough" attitude that mothers high on Authoritarian-Control take toward their children is an indication of their basic need to assume a more dominant, masculine role in the family. Many observers of the changing family relationships in American culture have commented on the increasing tendency for the mother to assume the directive role in the family with increasing lack of differentiation between the masculine and feminine roles in the family.

The Hostility-Rejection factor seems to be more closely related to psychopathological tendencies. Using the MMPI manual to define the typical abnormal behavior patterns measured by the scales, we may say that Hostility-Rejection is directly related to tendencies toward preoccupation

CHILD DEVELOPMENT

with bodily problems (Hs); suspiciousness, oversensitivity, and delusions of persecution (Pa); excessive worry, lack of confidence (Pt); bizarre or unusual thoughts or behavior (Sc); marked overproductivity in thought and action, disregard for social conventions (Ma); and anxiety as measured by the Taylor Anxiety Scale. Relations between Hostility-Rejection and the MMPI validity scales (F and K) are marked, indicating a lack of defensiveness and a "plus-getting" tendency to admit symptoms in women high on this PARI factor. Hostility-Rejection seems to correlate more significantly with the scales defining the psychotic triad on the MMPI (Pa, Pt, Ss) than with scales defining the neurotic triad (Hs, D, Hy).

Authoritarian-Control shows a significant relationship with only two of the MMPI "clinical" scales, while Hostility-Rejection relates significantly to six "clinical" scales. The relationship between parental attitudes and psychopathology predicted in hypothesis 5 seems to be borne out only for Hostile-Rejecting attitudes.

The Meaning of the Factors

We find that the Authoritarian-Control factor seems to be closely related to general authoritarian tendencies which in turn are related to the educational level of the mother. The few personality relationships found with Authoritarian-Control, i.e., high deference and low affiliation, tend to also be characteristics of the authoritarian personality as defined by Adorno *et al.* Authoritarian-Control, in itself, bears no strong relationship to psychopathology as measured by the MMPI. The strong relationship with masculine tendencies is one which might lead to some theoretical speculation about the role of sex-identification in parental attitudes. We might hypothesize that a woman who identifies with a strong father figure will tend to exert more control in her handling of children than the one who identifies with a weaker father or mother. All of the mother's tendencies to assume the male role may be channelized into assuming a dominant role in the home.

The Hostility-Rejection factor seems to be related to hostility and weak nurturing needs in the personality. In patients this hostility may be given expression in an externalized form, e.g., paranoid projections, or internalized in compulsive, hypochondriacal, or bizarre ideational complaints. The relationship between Hostility-Rejection and psychopathology, as measured by the MMPI, seems to be more pronounced than that of the Authoritarian-Control factor. Accompanying the hostility in factor B is a lack of self-esteem. These women cannot accept themselves or their families, but it would be difficult to judge which tendency has etiological priority. Hostility as a compensatory reaction to feelings of low self-esteem is one possibility.

The Democratic Attitudes factor bears little relationship to any of the areas of personality. The authors of the PARI seem to be justified in viewing these scales as having little potentiality for validity.

Schaefer (9) has suggested that a two-dimensional scheme for describing parental attitudes may be constructed using autonomy-control as one axis

MARVIN ZUCKERMAN and MARY OLTEAN

and hostility-love as the other axis. This scheme would yield four basic extremes of maternal attitudes and behavior using the four possible combinations of factors: (a) the punitive mother (high control, high hostility); (b) the rejecting mother (low control, high hostility); (c) the over-protecting mother (high control, low hostility); (d) the over-indulgent mother (low control, low hostility). Future research should be oriented toward sampling these two dimensions and four combinations of parental attitudes and behavior using a variety of techniques. One of the difficulties with the PARI form used here is its inadequate sampling of scales measuring Hostility-Rejection. This factor appears to be of major importance since it seems to be less related to formal education and more related to personality needs and adjustment.

SUMMARY

Correlations between three factors previously extracted from the Parental Attitude Research Instrument and authoritarian attitudes and personality dimensions were obtained in three groups: (a) 60 female psychiatric patients, mostly mothers; (b) 24 mothers of college students; (c) 88 unmarried student nurses.

1. The Authoritarian-Control factor of the PARI was significantly and positively correlated with the California F Scale of Authoritarianism in both the patient and student nurse samples. The other factors were not related to this scale.

2. A number of predicted relationships were found between certain personality needs measured by the Edwards Personal Preference Scale and the Hostility-Rejection and Authoritarian-Control factors in a group of mothers of college students. Most of the predictions were not borne out in the unmarried student nurse group.

3. A measure of self-acceptance correlated negatively and significantly with the Hostility-Rejection factor of the PARI and not with the other factors.

4. A high positive correlation between the Authoritarian-Control factor and masculine tendencies as measured by the Mf scale of the MMPI was found. Education showed a high negative correlation with Authoritarian-Control. Few positive correlations between Authoritarian-Control and clinical scales were found.

5. Hostility-Rejection showed a marked correlation with the clinical scales composing the "psychotic triad," Paranoia, Psychasthenia, and Schizophrenia. Strong positive correlations were also evident in relation to the Manic and Anxiety scales. A marked negative correlation with the K scale and a positive correlation with the F scale were obtained, indicating the influence of defensive and "plus getting" tendencies on this factor.

The results were interpreted as indicating some relationship between personality variables and attitudes toward child rearing, and offering some

CHILD DEVELOPMENT

evidence for the construct validity of the parental attitude factors. Schaefer's two-dimensional method of analysis of parental attitudes and behavior was suggested as an outline for future work.

REFERENCES

1. ADORNO, T. W., FRENKEL-BRUNSWICK, ELSE, LEVINSON, D., & SANFORD, R. N. *The authoritarian personality*. New York: Harper, 1950.
2. BASS, B. M. Authoritarianism or acquiescence? *J. abnorm. soc. Psychol.*, 1955, 51, 616-623.
3. BERGER, E. M. The relation between expressed acceptance of self and expressed acceptance of others. *J. abnorm. soc. Psychol.*, 1952, 47, 778-782.
4. BLOCK, J. Personality characteristics associated with fathers' attitudes toward child-rearing. *Child Developm.*, 1955, 26, 41-48.
5. EDWARDS, A. L. *Edwards Personal Preference Schedule, Manual*. New York: Psychological Corp., 1954.
6. HART, I. Maternal child-rearing practices and authoritarian ideology. *J. abnorm. soc. Psychol.*, 1957, 55, 232-237.
7. HATHAWAY, S. R., & MCKINLEY, J. C. *Minnesota Multiphasic Personality Inventory, Manual*. New York: Psychological Corp., 1951.
8. MARK, J. C. The attitudes of mothers of male schizophrenics toward child behavior. *J. abnorm. soc. Psychol.*, 1953, 48, 185-189.
9. SCHAEFER, E. S. Organization of maternal behavior and attitudes within a two-dimensional space: An application of Guttman's Radex theory. *Amer. Psychologist*, 1957, 12, 401. (Abstract)
10. SCHAEFER, E. S., & BELL, R. Q. Development of a parental attitude research instrument. *Child Developm.*, 1958, 29, 339-361.
11. SEARS, R. R., MACCOBY, ELEANOR, & LEWIN, H. *Patterns of child rearing*. Evanston, Ill., Row Peterson, 1957.
12. SHOEN, E. J. The assessment of parental attitudes in relation to child adjustment. *Genet. Psychol. Monogr.*, 1949, 39, 101-148.
13. WILLIS, R. H. Political and child-rearing attitudes in Sweden. *J. abnorm. soc. Psychol.*, 1956, 53, 74-77.
14. ZUCKERMAN, M., BAER, M., & MONASHKIN, I. Acceptance of self, parents and people in patients and normal. *J. clin. Psychol.*, 1956, 12, 327-332.
15. ZUCKERMAN, M., BARRETT-RIBBACK, BEATRICE, & MONASHKIN, I. Normative data and factor analysis on the parental attitude research instrument. *J. consult. Psychol.*, 1958, 22, 165-171.
16. ZUCKERMAN, M., & OLTEAN, MARY. The parental attitudes of mothers of schizophrenics. *J. consult. Psychol.*, 1958, 22, 307-310.

A CONTRIBUTION TO RESEARCH IN THE AREA OF THE MOTHER-CHILD RELATIONSHIP

HELEN FAIGIN ANTONOVSKY*

The Hebrew University

This paper is a description of an exploratory study in the area of the mother-child relationship.¹ Two general problems were investigated, one theoretical and the other methodological.

The theoretical problem was concerned with the testing of certain hypotheses about the relationship between child-rearing practices of mothers and certain behaviors of children just under two years of age, or what we have called the transition period from infancy to childhood.

The methodological problem involved the analysis of correspondences and discrepancies among three sets of data on mother behavior: structured interviews, observations, and unstructured interviews.

The theoretical orientation of the study, the choice of variables, and the hypotheses which were tested follow the general approach of Sears *et al.* (1, 2).

We chose three child behavior variables, four mother behavior variables, and formulated three general hypotheses about the relationship between these two sets of variables. The definitions of the variables and the hypotheses which were tested are given below.

CHILD BEHAVIOR VARIABLES

1. *Dependency* is defined as behavior whose goal is seeking help or attention. Every child is dependent upon someone (usually his mother) to satisfy his needs. How they are satisfied and the degree to which they are

* Bank Street, College of Education, New York City.

¹ The research was carried out as part of the work of the Lasker Mental Hygiene and Child Guidance Center of Hadassah in Jerusalem. The writer wishes to thank the members of the staff of the Lasker Center and its former directors, Dr. F. H. Stone and Dr. George Mohr, for their many helpful suggestions. Special thanks are rendered to Ruth Grushka, Schlomit Hoek, Hilde Rappaport, and Aliza Segall who helped in the collection and analysis of the data.

CHILD DEVELOPMENT

satisfied seems relevant to the frequency and intensity of dependent responses. We have used three measures of dependency responses:² (a) *seeking help*—a measure of dependency in which the child seeks help from the mother; (b) *affectional contact*—a measure of dependency in which the child seeks the mother's attention by making affectionate responses toward her; (c) *noninteractive play*—a measure of independence in which the child plays by himself, ignoring the mother's presence.

2. *Aggression* is defined in terms of behavior whose goal is causing physical or psychological pain. The frequency of occurrence and intensity of aggressive behavior is probably related to conditions under which such behavior is instigated, e.g., frustration, and the degree to which its occurrence is rewarded or punished. We have used one measure of aggression in this study, disobedience. Disobedience was measured in terms of the relative number of noncompliant responses to compliant responses on the part of the child to the mother's suggestions, initiations, or preventions of various activities.

3. *Initiative* is defined in terms of behavior which has as its goal achievement and the desire to do something independently. It probably involves some degree of creativity, doing something new on one's own. The frequency of occurrence of initiative behavior is probably related to conditions under which desire for independence and achievement are encouraged. Initiative was measured in terms of the child's suggesting and initiating new play activities.

MOTHER BEHAVIOR VARIABLES

1. *Affectional Contact* refers to the kinds of interaction in which mother and child participate to their mutual enjoyment and the kind of enjoyment they get out of being with one another. Relevant kinds of behavior are playing together, spontaneous display of affection, time spent together outside of routine caretaking activities.

2. *Expectations (Level of Demands)* refers to the kinds of demands which the mother makes of the child in terms of the standards of behavior and achievements which she sets for him at different age or developmental levels. The central focus is what the mother expects when. This behavior is relevant to the mother's general conception of childhood and orientation toward achievements. This behavior has both rewarding and punishing aspects for the child. It is psychologically punishing when the child does not meet the standards set for him and rewarding when he does.

3. *Restrictiveness* refers to the degree to which the mother circumscribes and limits the child's play and exploratory activities in terms of the dangers, physical and psychological, which the mother feels are inherent in the

² All measures of child behavior were made during an observation session in which both mother and child were present. The specific operational definitions of these behaviors are given in the section on methods and procedures.

HELEN FAIGIN ANTONOVSKY

environment. These dangers may be realistic or imaginary. Relevant behavior is the mother's supervision of the child and the rules and regulations which she applies to the child's behavior. Restrictiveness has the quality of inducing frustration but not permitting the child to learn responses for dealing with the frustration due to limitation of his opportunities for exploration.

4. *Punishment* refers to the mother's attempts to control the child's behavior by applying punishments to change or inhibit the behavior of the child which she considers wrong. The techniques of punishment vary but they have the common factor of being pain-inducing, physically or psychologically.

HYPOTHESES

Three general hypotheses were tested about the relationships which we expected to find between the variables of mother behavior and the variables of child behavior.

1. The frequency of *dependency* behavior on the part of the child is *negatively* related to the degree of *affectional contact* expressed by the mother and *positively* related to the degree of *demands, restrictiveness, and punishment*.

This hypothesis follows from conflict theory (2, p. 180), according to which we would expect the frequency of dependent responses to be greater where the child has conflicting expectancies about reward and nonreward or punishment for dependent behavior. Since all children as infants are rewarded for dependent behavior to some degree, the conflict should be greater under conditions of greater nonreward or punishment and, hence, the frequency of dependent responses greater.

With regard to our three measures of dependency we would thus expect: (a) the frequency of *help-seeking* and *affectional contact* responses on part of the child to be *negatively* related to *affectional contact* and *positively* related to *demands, restrictiveness, and punishment*; (b) the frequency of *noninteractive* responses to be *positively* related to *affectional contact* and *negatively* related to *demands, restrictiveness and punishment*, since non-interactive play is a measure of independence.

2. The frequency of *aggressive* behavior expressed toward the mother (disobedience) is *positively* related to the degree of *restrictiveness* and *negatively* related to the degree of *demands and punishment*.

This hypothesis follows from the frustration-aggression hypothesis, and the hypothesis about anxiety and avoidance. The frustration-aggression hypothesis leads us to expect that one of the most frequent responses to frustration is aggression. The anxiety-avoidance hypothesis leads us to expect that a response which has been punished in a given situation will tend not to occur again in that situation. Thus, the greater the frustration (restrictiveness), the greater the tendency for aggression to occur, and the

CHILD DEVELOPMENT

greater the discipline (demands and punishment), the greater the tendency for the response *not* to occur in the same or similar situation. In addition, the greater the degree of demands, the greater the opportunity to learn responses other than aggression to frustration.

3. The frequency of *initiative* behavior is *positively* related to the degree of *affectional contact* and *demands* and *negatively* related to the degree of *restrictiveness*.

This hypothesis follows from our assumptions in the definition of initiative: that it has components of independent behavior and orientation toward achievement. We expect, thus, that independence is positively related to affectional contact and achievement to demands. We further stated in our definition of restrictiveness that it tends to limit the child's opportunities for exploration and trying out new responses; therefore, we would expect restrictiveness to be negatively related to initiative.

METHODOLOGICAL PROBLEM

To get back to our methodological problem: A subject that has long bothered students of human behavior, but about which little systematic research has been done, is the problem of the relationship between interview and observation data. Some investigators (such as the classical behaviorists) have tended to dismiss verbal report almost completely as an invalid measure of behavior, whereas others (such as psychoanalysts) have used verbal report almost completely as a basis for the understanding of behavior. We do not wish to discuss the relative validity of verbal report and observation material as measures of behavior. Each, it seems, is valid in its own right depending upon the problem under investigation. Indeed, in much clinical work and in various types of psychological, sociological, and anthropological research, observation and interview material have been used to supplement each other in getting a better understanding of the problem at hand. Generally, when discrepancies are found between these sets of data, one or the other is usually dismissed on the basis of the judgments of the clinician or researcher.

Our problem here is to investigate the conditions under which one finds correspondences among sets of data on similar behavior gathered by different operations and those under which one finds discrepancies. We obtained three sets of data on similar aspects of mother behavior by using three different methods of data collection: structured interviews, observations, and unstructured interviews. In addition to finding out which set (or sets) of data give us better correlations between our two sets of variables, we were interested in testing hypotheses about the relations among these three sets of data themselves.

Our first hypothesis is self-evident. We expected to find closer correspondence between the two sets of interview data than between each of these sets of data and the observation data, since both involve more similar

HELEN FAIGIN ANTONOVSKY

operations than the collection of observation material. In addition, interviews are subject to the problems of verbal expression and recall.

Our second hypothesis is based on material from studies of ego-involvement and general clinical impressions. We expected that there would be greater discrepancies among the ratings based on the three sets of data for those mothers who were clinically judged to be "anxious" and "conflictful" than for mothers who were not judged "anxious." In addition, we expected these discrepancies to be greater in those particular areas of behavior in which they were judged to be particularly anxious or conflictful. If this is the case, discrepancies between interview and observation material could be useful as a measure of anxiety.

To sum up, the aims of this study were to test three hypotheses about the relationship between mother behavior and child behavior and two hypotheses about correspondences and discrepancies among three sets of data on mother behavior.

METHOD

Sample

The sample consisted of 9 mothers and their children (6 boys and 3 girls) ranging in age from 20 to 23 months who were in regular attendance at the neighborhood Infant Welfare Center. The basic criterion for sample choice, in addition to regular attendance at the Infant Welfare Center, was the age of the child. Another stipulation was that no referral of the mother-child pair had been made to the Lasker Center as a case of disturbance in the mother-child relationship. We were particularly interested in the nature of mother-child interaction in the period of transition from infancy to childhood in "normal" mother-child pairs. Seven of the children were first born, two were second children.⁸

It should be noted that all the mothers who participated in this study were interested in being "modern" mothers and wanted to know the latest psychological theories relating to how to bring up their children. It is to be emphasized that the sample is not considered representative of any wider population. Rather, the sample was used in an exploratory way as a first step in further formulation of hypotheses which would later be tested on a larger selected sample.

Procedure

Each mother was invited to come to the Lasker Center for three separate sessions, all of which took place within a period of 10 days.

The first session consisted of a structured interview which lasted approximately an hour in which the mothers were asked about their attitudes,

⁸ The original plan had been to have only single child families. The number of cases is small due to the fact that in the particular neighborhood from which the mother-child pairs were chosen, there were very few one child families. In most cases, the mother gives birth every year or year and a half.

CHILD DEVELOPMENT

values, and practices with reference to various areas of child rearing. The questions were not set and predetermined but the topics to be discussed were.

The second session consisted of a half-hour session in which the mother and child were observed in interaction with one another. The playroom was arranged in a standard way. Toys were chosen which would tend to induce interaction between mother and child. Mother and child behavior were recorded separately.

The third session consisted of an unstructured interview which lasted for an hour. The mother was told that she could discuss anything that she wished. Thus, in essence, the mother set the pace for this interview and chose the topics for discussion herself. This interview was similar to the usual clinic intake interview.

Rating scales were devised relating to the four mother behavior variables described in the previous section, affectional contact, demands, restrictiveness, and punishment. The scales were 7-point scales. Each mother was given a rating on each scale from each set of data. Thus, each mother obtained three sets of ratings on each behavior variable.⁴ The scale definitions were the same, but more detailed, than the definitions given in the previous section.

The ratings on each mother were made separately from each set of data by two independent judges. The reliability (Pearson r) of the ratings were as follows:

	<i>Structured Interview</i>	<i>Observation</i>	<i>Unstructured Interview</i>
Affectional Contact	.84	.81	.91
Demands97	.86	.63
Restrictiveness . .	.80	.89	.91
Punishment82		.83

Whenever the raters disagreed by more than one scale point, the ratings were discussed and the final rating which was used was a discussed rating. The reliabilities reported above are the correlations between the ratings of the two independent judges, and do not include these discussed ratings.

The material from the unstructured interviews was used, in addition, to make clinical judgments of each of the mothers on the four behavior variables described above. These judgments were made by four separate judges: two psychiatrists, a psychiatric social worker, and a psychologist. These judgments were made independently, then discussed, and a final, agreed upon judgment was reached. These clinical judgments were made mainly in terms of the anxiety and conflict which characterized the interaction between mother and child in the four areas of mother behavior under investigation.

The observation protocols were divided into behavior units. Behavior units were defined as the smallest unit of molar behavior which indicates

⁴ Exceptions: no ratings on punishment could be made from the observation data and three mothers could not be rated on restrictiveness from the unstructured interview data.

HELEN FAIGIN ANTONOVSKY

action, has a beginning and end, and is separable from other acts. For example, "takes cubes with both hands, puts them back, throws ball, smiles, gets ball," would be counted as five separate behavior units. All marking was done by the project director who had previous training in this area. The first few protocols were also marked off in units, in addition, by other members of the research team. The agreement on the three protocols which were marked by more than one person was 85 per cent, using Festinger's formula of number of agreements over the number of disagreements.

The data from the observation protocols on child behavior were categorized in terms of the following categories:

Dependency

Help—seeking mother's help in carrying out some action.⁵

Affectional Contact—any signs of affection expressed by child to mother, such as smiling, kissing, hugging, playing near.

Noninteractive Play—child plays by himself, ignores mother's presence.

Aggression (the ratio Noncompliance / Compliance in relation to responses of mother which could evoke one of these two responses)

Noncompliance—ignoring or actively refusing to carry out mother's suggestions; doing what he has been told not to do.

Compliance—doing or carrying out mother's suggestions on activities and play; acting in accordance with mother's restriction of activities.

Initiative—suggestions by child to mother of play activities; initiating new activities.

The reliability of categorizing the data from the protocols was calculated using the Festinger method. The over-all reliability thus obtained was 87 per cent. Using these categories, the following correlations were obtained as measures of reliability between the two observers:

	<i>r</i>
Help74
Noninteractive Play95
Affectional Contact75
Noncompliance93
Compliance91
Initiative94

The scores used for the children were percentage scores of the number of units in a particular behavior category over the total number of units in the protocol.⁶ Thus, if a child gave 30 initiative responses out of a total of 150 responses, his score on initiative was 20 per cent. In the case of the aggression score a combination score was used: The ratio of noncompliant

⁵ The room was arranged in such a way that the average child of this age could do or reach everything by himself. Thus, most cases of help seeking were not due to the child's inability to do something by himself, but by a presumed wish to get the mother's help or attention.

⁶ Where the percentage scores of the observers were different, the median between the two scores was used. Thus, if Observer A had 20 per cent and Observer B, 24 per cent, the final score was 22 per cent.

CHILD DEVELOPMENT

responses over the compliant responses using the total number of suggestion and restriction responses of the mother as a base rather than the total number of units in the protocol.

RESULTS

Mother-Child Relationships

In this section we present the findings concerning the relationships between mother behavior and child behavior. Our first hypothesis dealt with the relationship between dependency behavior on the part of the child and the four areas of mother behavior under investigation. Each hypothesis is restated below followed by the findings.

1. The frequency of occurrence of dependent behavior on the part of the child is negatively related to the degree of affectional contact expressed by the mother toward the child and positively related to the degree of demands, restrictiveness, and punishment expressed by the mother toward the child.

a. The frequency of the help-seeking responses on the part of the child is negatively related to the degree of affectional contact expressed by the mother toward the child and positively related to the degree of demands, restrictiveness, and punishment expressed by the mother toward the child.

The results of the test of this hypothesis are presented in Table 1. Although not significant, they tend to fall in the predicted direction. It should be noted also that on the whole the data from the structured interview give us the clearest trends.

TABLE I
MOTHER BEHAVIOR AND CHILD BEHAVIOR—SEEKING HELP

<i>Mother Behavior</i>	<i>Structured Interview</i>	<i>Observation</i>	<i>Unstructured Interview</i>
Affectional Contact	-.50	-.08	+.04
Demands	+.31	+.12	+.50
Restrictiveness	+.40	+.10	..
Punishment	+.37	..	+.03

NOTE.—All correlations in this table and subsequent tables presenting the data on the relationship between mother behavior and child behavior are rank order correlations.

b. The frequency of occurrence of affectional contact responses on the part of the child is negatively related to affectional contact on the part of the mother and positively related to demands, restrictiveness, and punishment.

The results, presented in Table 2, do not support the hypothesis. Although the correlations are low, the trends are in the same direction as those of noninteractive play (see Table 3). Affectional contact was considered a measure of dependency, since it seemed that expression of affection on the part of the child was a means of seeking attention and affection from the

HELEN FAIGIN ANTONOVSKY

TABLE 2

MOTHER BEHAVIOR AND CHILD BEHAVIOR—AFFECTIONAL CONTACT

<i>Mother Behavior</i>	<i>Structured Interview</i>	<i>Observation</i>	<i>Unstructured Interview</i>
Affectional Contact	+.29	+.52	+.60
Demands	-.15	-.38	-.18
Restrictiveness	+.21	-.02	..
Punishment	-.24	..	-.07

mother. The results do not support this type of reasoning. Perhaps the expression of affection by the child is a special type of dependency response which requires expression of affection on the part of the mother in order for the child to learn this kind of response.

c. The frequency of occurrence of noninteractive play (independence) on the part of the child is positively related to the degree of affectional contact expressed by the mother and negatively related to the degree of demands, restrictiveness and punishment.

The results of the test of this hypothesis are presented in Table 3. They tend to support the hypothesis except with regard to the expected negative relationships between noninteractive play and restrictiveness. It is not clear why this should be the case. Perhaps, the observation situation as it was set up provided the child who is constantly under his mother's supervision one of the few opportunities he has to play by himself. Or, alternatively, it may be that the observation session itself was a restricting situation in that the mother and child were confined together in one room. Perhaps, under these kinds of conditions the child has been able to build up the habit of play by himself; that is, the child can play by himself within the confines of a supervised setting, whereas we would expect less noninteractive play among the children whose mothers are highly restrictive in a less confined

TABLE 3

MOTHER BEHAVIOR AND CHILD BEHAVIOR—NONINTERACTIVE PLAY

<i>Mother Behavior</i>	<i>Structured Interview</i>	<i>Observation</i>	<i>Unstructured Interview</i>
Affectional Contact	+.59	+.85**	+.78**
Demands	-.51	-.86**	-.70*
Restrictiveness	+.43	+.74*	..
Punishment	-.71*	..	+.21

* Significant at .05 level.

** Significant at .01 level.

CHILD DEVELOPMENT

situation. It should be noted that all three sets of data tend to show the same trend. This time the correlations using the structured interview tend to be lower than those using the other two sets of data.

2. The frequency of disobedience on the part of the child as measured in terms of the relative noncompliance to compliance to mother's suggestions, interference with, and prevention of activity on the part of the child is positively related to restrictiveness and negatively related to demands and punishment.

The results of the test of this hypothesis, contained in Table 4, although not significant, tend to support the hypothesis. All the correlations are in the predicted direction. These data are in conflict with those reported by Sears *et al.* (1, 2). In their groups of kindergarten children they found that punishment tended to increase the amount of aggression and restrictiveness and demands tended to decrease the amount of aggression shown by the children as reported by their mothers. The differences here may be due to a number of factors, age of the children and the way in which both aggression and restrictions and demands were measured.

TABLE 4

MOTHER BEHAVIOR AND CHILD BEHAVIOR—NONCOMPLIANCE/COMPLIANCE

<i>Mother Behavior</i>	<i>Structured Interview</i>	<i>Observation</i>	<i>Unstructured Interview</i>
Demands	-.49	-.23	-.07
Restrictiveness	+.58	+.27	..
Punishment	-.05	..	-.14

3. The frequency of occurrence of initiative responses on the part of the child is positively related to the degree of affectional contact and demands expressed by the mother toward the child and negatively related to the degree to which the mother restricts the child. The results of the test of this hypothesis are presented in Table 5; they tend to support the hypothesis

TABLE 5

MOTHER BEHAVIOR AND CHILD BEHAVIOR—INITIATIVE

<i>Mother Behavior</i>	<i>Structured Interview</i>	<i>Observation</i>	<i>Unstructured Interview</i>
Affectional Contact and Demands	+.50	+.63	-.22
Restrictiveness	-.65*	-.32	..

* Significant at .05 level.

HELEN FAIGIN ANTONOVSKY

except for the data from the unstructured interview. This problem will be taken up in a later section when we discuss the analysis of the correspondences and discrepancies among the three sets of data.

Correspondences and Discrepancies Among the Three Sets of Data on Mother Behavior

The data were analyzed to find out if any one set of data gives us better correlations in terms of our predictions of mother-child relationships than any other set. From the tables presented in the previous section it may be seen that out of a total of 12 correlations predicted according to our hypotheses the highest were obtained from the structured interview data in seven cases, three out of a possible nine⁷ from the observation data, and two out of a possible nine⁸ from the unstructured interview data. Thus, we see that we get the best over-all set of correlations using the structured interview data. In addition, it may be noted that in three cases the unstructured interview data give us correlations in the opposite direction from that predicted, whereas the other two sets of data give us correlations in the predicted direction in these same cases. When we check the correlations which did not support our predictions, we find that out of five possible relationships the structured interview gives us the highest correlations two times, the observation data in two out of four cases, and the unstructured interview data only one out of three times. It appears, then, that the ratings from the structured interview data and observation data tend to give us better correlations with our child behavior measures than do the ratings based on the unstructured interview.⁹

According to our hypothesis about the three sets of data obtained on mother behavior, we expected to find closer correspondence between the two sets of interview material than between each set of interview material and the observation data. The correlations (Pearson r) among the three sets of ratings are presented in Table 6.

These data do not support our hypothesis. In the case of demands and restrictiveness we do get higher correlations between the two sets of interview data than between each set and the observation. However, this is not the case in the ratings on affectional contact where the best correlation is between the observation ratings and unstructured interview ratings. In the case of punishment we find a negative correlation between the two sets of interview data.

⁷ The observation data do not include a measure of punishment so that a complete set of 12 relationships could not be calculated.

⁸ There is no measure of restrictiveness for the unstructured interview since three mothers could not be rated on this variable due to lack of data.

⁹ The intercorrelations among the variables of mother behavior were lower, in general, for the structured interview data than for either the unstructured interview or the observation data. Among the child behavior measures the only correlation of any significance was that between noncompliance and initiative, perhaps due to the way in which non-compliance was defined.

CHILD DEVELOPMENT

TABLE 6

INTERCORRELATIONS AMONG THREE SETS OF DATA ON
MOTHER BEHAVIOR

	<i>Affectional Contact</i>	<i>Demands</i>	<i>Restrictiveness</i>	<i>Punishment</i>
Structured Interview and Unstructured Interview	+ .40	+ .64	+ .38*	-.33
Structured Interview and Observation	+ .35	+ .42	-.20	..†
Unstructured Interview and Observation	+ .64	+ .30	-.08*	..†

* This correlation is based on only six cases since the other three could not be rated on restrictiveness from the data in the unstructured interview.

† No ratings on punishment could be made from the observation data.

What does all this mean? These results indicate that there is no closer relationship between the sets of interview data collected in different ways than between interviews and observations. In addition, if we use as a measure of validity of mother behavior data correlations with child behavior in the predicted direction, we should find closer correspondence in this case between the ratings of the structured interviews and observation than between each of these sets of ratings and the unstructured interview ratings. Again, this is not the case. The reliabilities of the rating judgments on each set of data were high so that lack of reliability does not seem to account for the lack of correspondence among the sets of data.

All this brings us to a consideration of our next hypothesis. Perhaps the general lack of correspondence among the three sets of data is related to conditions operating on the individuals from whom the data were collected.

We found in looking over the individual ratings that in certain cases there were close correspondences among the three sets of ratings and in other cases there were wide discrepancies. These correspondences and discrepancies are presented in Table 7.

It will be noted from Table 7 that mother 9 has the highest over-all discrepancy score, that mother 3's discrepancy score is highest in the area of restrictiveness, that mother 2's highest discrepancy score is in the area of demands, etc.

Our hypothesis states that we would expect higher discrepancy scores among mothers who were judged anxious and conflictful and, in addition, these higher discrepancy scores would show up especially on those behavior variables about which the mother was judged anxious and conflictful. Our measure of anxiety and conflict was obtained from clinical judgments based on the data from the unstructured interview. The clinical assessments were made only on the basis of the unstructured interview, and none of these judges saw the ratings from the structured interview or observation.

HELEN FAIGIN ANTONOVSKY

TABLE 7

DISCREPANCY SCORES FOR THE THREE SETS OF DATA*

<i>Mother</i>	<i>Affectional Contact</i>	<i>Demands</i>	<i>Restrictiveness†</i>
1	2	2	2
2	2	6	..
3	4	4	8
4	4	0	..
5	4	6	6
6	0	2	2
7	4	2	..
8	2	2	4
9	8	10	6

* These scores are based on differences between the ratings on each set of data. For example, if a given mother were scored in the area of affectional contact as follows: structured interview 5, unstructured interview 3, observation 4, her total discrepancy score would be 4. Punishment is not included since no rating could be made from the observation data.

† No rating could be made on restrictiveness from the unstructured interview in these cases.

The data do not support the hypothesis about the relationship between anxiety and conflict and the discrepancy scores, neither over-all nor for any given area. This may mean that there is no relationship between anxiety and conflict in a given area of behavior and the tendency either to behave inconsistently or to report inaccurately on this behavior, or our measure of anxiety and conflict, the clinical assessments, may not be a good measure of this phenomenon.

In a study by Smith (3) "anxious" mothers were divided into two groups. In the first group were "anxious" mothers whose report of their behavior in an interview was inaccurate when compared with their behavior as observed. For these mothers anxiety was considered to be a basic personality trait, a general over-all anxiety. In the second group were mothers who were anxious only in the sense that they were upset or "anxious" about their behavior in a given area in which they felt they were not doing "what they should do." The mothers in the second group reported very accurately their behavior which they felt deviated from "proper" behavior. Our measure of anxiety, a general clinical judgment, may have lumped these two types of mothers together, those giving accurate reports and those giving inaccurate reports, into one group, and for this reason we find no relationship between anxiety and discrepancy. It is felt that the test of this hypothesis made here is not adequate. A more adequate measure of anxiety is needed. General clinical impressions and studies of recall and ego involvement all tend to indicate that this is an area to which more systematic attention might be paid.

CHILD DEVELOPMENT

SUMMARY AND CONCLUSIONS

The study reported here was an exploratory investigation in which two sets of hypotheses were tested. One set of hypotheses was concerned with the relationship between mother behavior and child behavior, the other, with the relationship among three sets of data on mother behavior obtained by use of different methods of data collection. The sample consisted of nine mothers and nine children aged 20 to 23 months of age.

Our empirical findings may be summarized as follows in terms of the hypotheses which we tested:

Hypothesis 1

a. The frequency of help-seeking and affectional contact responses on the part of the child is negatively related to the degree of affectional contact expressed by the mother and positively related to the degree of demands, restrictiveness, and punishment expressed by the mother toward the child. Our data tended to support this hypothesis about help-seeking responses although none of the correlations was significant. The data do not support the hypothesis about affectional contact responses.

b. The frequency of occurrence of noninteractive play responses on the part of the child is positively related to the degree of affectional contact expressed by the mother and negatively related to the degree of demands, restrictiveness, and punishment. Our data tended to support this hypothesis except that noninteractive play was positively related to restrictiveness, not negatively related as predicted.

Hypothesis 2

The frequency of occurrence of aggressive behavior (disobedience) on the part of the child is positively related to the degree of restrictiveness and negatively related to the degree of demands and punishment. Our data tended to support this hypothesis.

Hypothesis 3

The frequency of occurrence of initiative responses on the part of the child is positively related to the degree of both affectional contact and demands expressed by the mother toward the child and negatively related to the degree to which the mother restricts the child. The data tended to support this hypothesis.

These results thus support the general assumption of a predictable influence of mother behavior on child behavior. Our findings tend to support some previous findings relating to conflict theory, the frustration-aggression hypothesis, and the anxiety-aggression hypothesis, but not others. Our findings on initiative take us a further step in understanding conditions influencing achievement orientation. Although our sample was very small and the correlations for the most part not statistically significant, our findings indicate that certain learning conditions set by the mother in her child-

HELEN FAIGIN ANTONOVSKY

rearing patterns influence the development of certain behavior patterns and motivations in her child. Further investigation of the nature of restrictiveness and further definition of different types of dependency responses must be made, especially the meanings of expression of affection by the child. Further testing of these hypotheses must be made on larger, carefully selected samples before any real conclusions may be reached.

Two hypotheses about the correspondences and discrepancies among structured interviews, observations, and unstructured interviews were tested.

According to our first hypothesis we expected to find closer correspondences between the ratings from the two sets of interview data than between each set and the ratings from the observation data. Our findings do not support this hypothesis. There are low and even negative correlations among the ratings from the three sets of data.

According to our second hypothesis, we expected that there would be higher discrepancy scores among mothers who were judged anxious and conflictful than among mothers judged as nonanxious. Our findings do not support this hypothesis.

It is felt that this problem necessitates further investigation along similar lines but with a clearer definition of what is meant by the terms "anxiety" and "conflict."

REFERENCES

1. SEARS, R. R., MACCOBY, ELEANOR E., & LEVIN, H. *Patterns of child rearing*. Evanston, Ill.: Row, Peterson, 1957.
2. SEARS, R. R., WHITING, J. W. M., NOWLIS, V., & SEARS, PAULINE S. Some child-rearing antecedents of aggression and dependency in young children. *Genet. Psychol. Monogr.*, 1953, 47, 135-234.
3. SMITH, HENRIETTA T. A comparison of interview and observation measures of mother behavior. *J. abnorm. soc. Psychol.*, 1958, 57, 278-282.
4. STONE, F. H. A critical review of a current program of research into mother-child relationship. In G. Caplan (Ed.), *Emotional problems of early childhood*. New York: Basic Books, 1955. Pp. 95-116.

RACIAL COMPARISONS AND ADDITIONAL NORMATIVE DATA ON THE CHILDREN'S MANIFEST ANXIETY SCALE

DAVID S. PALERMO*¹

Southern Illinois University

Recently, Castaneda, McCandless, and Palermo (1) reported an adaptation of the Taylor manifest anxiety scale (6) for use with fourth, fifth, and sixth grade children. In addition, they report some relationships between the scale and complex learning tasks (2, 5), school achievement and intelligence (3), and social status (4).

The present paper reports some additional normative and reliability data relevant to the scale including a Negro, as well as a white, population. The reliability data is based on a one-month interval between tests rather than the one-week interval used in the previous study (1).

METHOD

Subjects. The children's form of the manifest anxiety scale (CMAS) (1) was given to all fourth, fifth, and sixth grade children in the seven Carbondale public grade schools and to these same grades in one public school of Negro students in a neighboring town. Included in the group were 61 Negro boys, 75 Negro girls, 207 white boys, and 187 white girls. Only Ss in the Carbondale schools were given the second administration

* Institute of Child Development and Welfare, University of Minnesota, Minneapolis 14.

¹ This study was supported in part by the Graduate School of Southern Illinois University. The author wishes to express appreciation to Mr. Basil Sherlock and Mr. Ronald Brown for their valuable work in helping to secure and process the data on which this paper is based. The author also wishes to express his appreciation to Mr. Norman Moore, Elementary School Supervisor in Carbondale, and Mr. William Carruthers, Superintendent of the Murphysboro Public Schools for their cooperation and to all the principals and teachers involved for their help in making this study possible.

CHILD DEVELOPMENT

of the test. Due to this fact and to absences, the reliability correlations are based on 44 Negro boys, 55 Negro girls, 196 white boys, and 175 white girls.

Procedure. Each teacher was given enough copies of the test for each member of the class, and the teacher administered the test on a group basis. In addition to the instructions on the test, the teacher was instructed to indicate that she was not going to look at the papers and to answer any questions the children asked about the items by telling them to answer "truthfully." The test was re-administered in the same manner one month later.

TABLE I
ANXIETY SCALE MEANS AND SD'S FOR THE DIFFERENT GRADE, SEX,
AND RACIAL GROUPS SEPARATELY

	GRADE								
	FOURTH			FIFTH			SIXTH		
	N	M	SD	N	M	SD	N	M	SD
<i>Negro</i>									
Boys	23	22.91	5.81	21	22.52	5.57	17	16.59	7.00
Girls	24	23.29	7.39	27	23.59	5.67	24	20.75	5.39
<i>White</i>									
Boys	79	17.04	8.07	65	14.23	8.26	63	15.17	7.64
Girls	70	17.93	8.12	54	17.37	7.08	63	16.56	6.45

RESULTS

Table I indicates the means, standard deviations, and number of Ss in each racial, sex, and grade group for the 42 anxiety items of the scale only.

It will be noted that the mean anxiety scores for the Negro Ss are higher than those for the white Ss at all grade levels and for both sexes. Statistical analysis of the over-all differences between the Negroes and the whites resulted in a *t* value of 7.36 which with 528 degrees of freedom is significant at well beyond the .001 level of confidence.

As was the case with the sample reported earlier (1) on this part of the scale, and on the Taylor scale (6), the girls tend to score higher than the boys. In this case, both Negro and white girls tend to score higher than the males in the respective racial groups at each grade level. A statistical test of the differences in scores of the two sexes yielded a *t* value of 2.95 which with 528 degrees of freedom is significant at the .01 level of confidence.

It may also be noted in Table I that there is some indication of an inverse relationship between grade and anxiety. A simple analysis of variance of these data resulted in an *F* value of 3.43 which with 2 and 527 degrees of freedom is significant at the .05 level of confidence.

DAVID S. PALERMO

TABLE 2

L-SCALE MEANS AND SD'S FOR THE DIFFERENT GRADE, SEX,
AND RACIAL GROUPS SEPARATELY

	G R A D E								
	F O U R T H			F I F T H			S I X T H		
	N	M	SD	N	M	SD	N	M	SD
<i>Negro</i>									
Boys	23	5.09	1.93	21	4.14	1.64	17	4.94	2.13
Girls	24	5.50	1.63	27	4.40	1.85	24	5.25	1.92
<i>White</i>									
Boys	79	4.66	1.90	65	4.08	2.09	63	2.98	2.20
Girls	70	4.73	2.17	54	4.85	2.23	63	3.51	2.15

Table 2 indicates the means, standard deviations and number of Ss in each racial, sex, and grade group for the 11 L-scale items of the scale.

With the exception of the fifth grade girls, all of the Negro groups have higher scores on this part of the scale, as well as on the anxiety items. Analysis of these data resulted in a t value of 3.58 which with 528 degrees of freedom is significant at well beyond the .01 level of confidence.

The girls tend to score higher than the boys on this part of the test too, a finding similar to that previously reported. A test of this difference resulted in a t value of 2.26 which with 528 degrees of freedom is significant at the .05 level of confidence.

TABLE 3

ANXIETY SCALE RETEST CORRELATIONS FOR THE DIFFERENT
GRADE, SEX, AND RACIAL GROUPS SEPARATELY

	G R A D E					
	F O U R T H		F I F T H		S I X T H	
	N	r	N	r	N	r
<i>Negro</i>						
Boys	18	.63	16	.80	10	.86
Girls	17	.91	21	.84	17	.59
<i>White</i>						
Boys	76	.89	61	.86	59	.75
Girls	66	.61	50	.77	59	.74

NOTE.—All correlations are significant at the .01 level.

CHILD DEVELOPMENT

As with the A-scale, there appears to be an inverse relationship between L-scores and grade. A simple analysis of variance performed on these data resulted in an F value of 13.94 which with 2 and 527 degrees of freedom is significant at the .01 level indicating a significant difference between grades.

In general, the A-scale scores are lower and the L-scale scores are higher for this population than for the one originally reported (1).

Table 3 presents the test retest reliability Pearson product-moment correlation coefficients of the A-scale scores broken down into grade, sex, and racial groups. As may be seen from the table, all of the correlations are significant beyond the .01 level of confidence indicating that the relative position of the S in his group tended to remain constant over the one-month period. The correlations, as might be expected, are somewhat lower than those reported in the original study when only one week separated the two testing periods.

TABLE 4
L-SCALE RETEST CORRELATIONS FOR THE DIFFERENT
GRADE, SEX, AND RACIAL GROUPS SEPARATELY

	GRADE					
	FOURTH		FIFTH		SIXTH	
	<i>N</i>	<i>r</i>	<i>N</i>	<i>r</i>	<i>N</i>	<i>r</i>
<i>Negro</i>						
Boys	18	.77*	16	.80*	10	.51
Girls	17	-.01	21	.27	17	.78*
<i>White</i>						
Boys	76	.53*	61	.65*	59	.54*
Girls	66	.98*	50	.39*	59	.68*

* Significant at the .01 level.

Table 4 presents the test retest reliability Pearson product-moment correlation coefficients of the L-scale scores broken down into grade, sex, and racial groups. As was the case in the original sample the reliabilities obtained for the white S s, although somewhat lower than the anxiety scale, are all significant at the .01 level of confidence. However, in this sample it may be noted that three of the Negro subgroup reliabilities do not reach statistically significant levels indicating less consistency from one testing to the next for the Negro population.

Table 5 presents the Pearson product-moment correlations between the two parts of the total test. The correlations indicate a tendency for the two scales to be inversely related although only four of the 12 correlations reach the .05 or greater level of significance.

DAVID S. PALERMO

TABLE 5

CORRELATIONS BETWEEN THE ANXIETY SCALE AND THE L-SCALE FOR THE DIFFERENT GRADE, SEX, AND RACIAL GROUPS SEPARATELY

	GRADE					
	FOURTH		FIFTH		SIXTH	
	N	r	N	r	N	r
<i>Negro</i>						
Boys	23	.08	21	-.14	17	-.08
Girls	24	-.04	27	-.44*	24	-.41*
<i>White</i>						
Boys	79	.15	65	-.33**	63	-.29*
Girls	70	-.11	54	-.12	63	-.18

* Significant at the .05 level.

** Significant at the .01 level.

SUMMARY

The present study reports some descriptive statistics obtained with the CMAS for a population of Negro and white fourth, fifth, and sixth grade children.

The mean scores for the white Ss were found to be lower on the A-scale and higher on the L-scale when compared with the group originally tested. One-month retest reliabilities were only slightly lower than those obtained with a one-week interval in the group originally tested.

The Negro group was found to score significantly higher than the white group on both parts of the test.

REFERENCES

1. CASTANEDA, A., McCANDLESS, B. R., & PALERMO, D. S. The children's form of the manifest anxiety scale. *Child Developm.*, 1956, 27, 317-326.
2. CASTANEDA, A., PALERMO, D. S., & McCANDLESS, B. R. Complex learning and performance as a function of anxiety in children and task difficulty. *Child Developm.*, 1956, 27, 327-332.
3. McCANDLESS, B. R., & CASTANEDA, A. Anxiety in children, school achievement, and intelligence. *Child Developm.*, 1956, 27, 379-382.
4. McCANDLESS, B. R., CASTANEDA, A., & PALERMO, D. S. Anxiety in children and social status. *Child Developm.*, 1956, 27, 385-391.
5. PALERMO, D. S., CASTANEDA, A., & McCANDLESS, B. R. The relationship of anxiety in children to performance in a complex learning task. *Child Developm.*, 1956, 27, 333-337.
6. TAYLOR, JANET A. A personality scale of manifest anxiety. *J. abnorm. soc. Psychol.*, 1953, 48, 285-290.

CHILDREN'S CONCEPTS OF JUSTICE: A COMPARISON WITH THE PIAGET DATA

DOLORS DURKIN*¹

University of California

This is an investigation of children's concepts of justice, in particular, of justice regarding one's person. It is specifically designed to examine developmental trends, with age, in these concepts, and to compare the trends with those described by Piaget in *The Moral Judgment of the Child* (9). It also examines the relationship between the intelligence of children and their justice-concepts.

In the Piaget investigation 167 subjects, ranging in age from 6 to 12 years, were questioned about what should be done "if someone punches you." They generally proposed two quite different solutions. Younger subjects favored reporting to an authority person; older subjects, a return of the aggression. Piaget saw in this trend of responses evidence for his theory concerning "the existence of a sort of law of evolution in the moral development of the child" (9, p. 225).

According to Piaget's exposition of the theory, this law is of psychological origin being rooted in child-adult social relationships. For the child of about 7 or 8 the relationship is necessarily one of unilateral respect, hence, of constraint. And, Piaget maintains, the fruit of this constraint in moral-judgment development is twofold: (a) an ethic of heteronomy wherein the good is defined by submission to adult authority, and (b) moral realism, "the tendency . . . to regard duty and the value attaching to it as self-subsistent and independent of the mind, as imposing itself regardless of

* School of Education, University of California, Berkeley 4.

¹ The author is especially indebted to Dr. J. Harlan Shores of the University of Illinois for the advice and guidance given in completing the doctoral dissertation upon which this article is based.

CHILD DEVELOPMENT

the circumstances in which the individual may find himself" (9, p. 106). At this first stage, then, justice is defined by submission to adult authority and, further, moral judgments are made without concern for the particular circumstances of the situation being judged.

Among children ranging in age from about 8 to 11, the child-adult relationship changes; it becomes one of mutual respect because "as the child grows up, his relations with adults approximate to equality . . ." (9, p. 84). Now, an ethic of autonomy or cooperation begins to make itself felt, and it is at this second stage that the equalitarian notion of justice emerges. Here, maintains Piaget, "reciprocity stands so high in the eyes of the child that he will apply it even where to us it seems to border on crude vengeance" (9, p. 215).

Later, in a period which begins at approximately 11 to 12 years, a new attitude develops parallel with the equalitarian notion of justice. It is an attitude Piaget calls "equity" which is "the development of equalitarianism in the direction of relativity" (9, p. 310). At this stage, then, moral judgments are made on the basis of the particular circumstances of a situation and, too, justice continues to be defined in terms of reciprocity.

The theory of Piaget and the findings he puts forth as substantiating the theory suggest the following problem for consideration: If subjects different in nationality and economic status from those included in the Piaget study, but similar in terms of chronological age, were questioned about an act of physical aggression, would their responses reflect stages in moral judgment identical to those outlined by Piaget? Or has he, on the basis of interviews with children "from the poorer parts of Geneva (Switzerland)" made too sweeping generalizations about children in general? Studies evaluating other aspects of Piaget's work (1, 2, 3, 4, 5, 6, 7, 8) indicate that he has.

A second problem, suggested by Piaget's emphasis on chronological age, concerns the function of intelligence in moral-judgment development. Might it not be that intelligence, rather than chronological age, is the significant factor in the development of both equalitarianism and equity?

The present study considers both of these problems.

PROCEDURE

School-Community Setting

The locale of the present study was a midwestern community with a population of about one thousand, all of whom were white and, with the exception of three Catholic families, also Protestant. The primary income-source for the community was farming. Three large factories were other important sources of income.

The school from which subjects were chosen was of the community-consolidated type. It was the only elementary school in the area and included grades 1 through 8. The total enrollment was 271.

DOLORES DURKIN

Subjects

The subjects were 101 boys and girls in grades 2, 5, and 8. All of the children in each of the grades chosen were included. These particular grade-levels were selected because in them it would be most likely to find chronological ages of 7, 10, and 13, the age-groups that correspond to what Piaget designates as "three great periods in the development of the sense of justice in the child" (9, p. 314). The actual mean ages of the three groups were 7.8, 10.9, and 13.9.

Since Piaget described the background of his subjects only in terms of their being poor, subjects in the present study were also classified in terms of economic status. The first-grade teacher in the school rated each subject as being "poor," "average," or "rich." She was chosen as the person most likely to have such information on the bases of life-long residence in the community, acquaintance with all of its families, and 35 years of teaching experience in this particular school. For the purpose of verification, her ratings were later compared with those independently given by the school clerk, also a resident of the community. In three instances their judgments showed classification differences; in each of these they decided together which of the ratings was more accurate and realistic.

Examination of their judgments showed that the economic rating "average" had been given to 75.0 per cent of the second graders, 68.4 per cent of the fifth graders, and 80.0 per cent of the eighth graders. Ratings of the remaining subjects were fairly equally divided between "rich" and "poor." It is on the basis of the total distribution of ratings that subjects in the present investigation are assumed to be different from Piaget's in terms of economic status.

Instrument

Following Piaget's procedure, an attempt was made to arrive at children's concepts of justice regarding one's person through their responses to questions about story-situations which depict possible violations of justice.

The following story-situation and questions were presented to male subjects:

- (A) One day when they were out at recess Bennett hit Van.
What should Van do?
Why?

To those subjects who, in their responses to question (A), appeared to subscribe to an equalitarian or "eye for an eye" concept of justice, the following situation and questions were also presented:

- (B) What if Van hit Bennett back and gave him a push besides?
What would you think about that?
Why?

CHILD DEVELOPMENT

Situation (B) was included for the purpose of evaluating Piaget's finding that children who accept an "eye for an eye" concept of justice do not accept "a sort of arbitrary punishment whose content bears no relation to the punishable act" (9, p. 299).

Story-situations and questions identical to (A) and (B), but involving girls ("Adla" and "Hannah"), were presented to female subjects.

Interviews

Subjects' oral responses to questions about the story-situations were obtained through tape-recorded interviews. For each interview only the investigator, acting as interviewer, and one subject were present.

At the beginning of each interview the investigator explained to the subject that she was interested in finding out what children really thought about certain kinds of things. She promised that nothing of what he said would be told to parents, principal, or teacher; and she also explained that he, as a respondent, would remain anonymous.

Because of the intentionally brief, almost skeleton-like form of the story-situations, it was expected that at least some of the subjects would ask questions in response to those of the interviewer. Certainly to be expected, for example, were such questions as "How old is he?" or "Did he do it on purpose?" or "Did he ever do that before?" If and when such questions were posed, the interviewer's standard response was, "Would that make a difference?" and "Why?"

Administration of Intelligence Test

Once all of the interviews were completed, the investigator administered the Kuhlmann-Anderson Intelligence Test to 100 of the subjects.² IQ scores in grade 2 ranged from 92 to 114. The mean score was 103.4; the standard deviation, 5.1. In grade 5, scores ranged from 69 to 122. The mean was 101.6, and the standard deviation was 13.2. For grade 8, scores ranged from 77 to 148. The mean was 103.4; the standard deviation, 11.4.

FINDINGS

Responses and Grade Level

Kinds of responses given to the problem of what a child should do if another child hits him, and the number of subjects at each grade level giving these responses, are summarized in Table 1.

To test the hypothesis that, within each grade, the different kinds of responses were chance occurrences, a chi square test was used. For grade 8, the hypothesis can be rejected with confidence at the 5 per cent level, the criterion of significance used throughout this paper. For grades 2 and 5, however, the hypothesis cannot be rejected ($.05 < p < .10$; and $.50 < p < .70$, respectively).

² One second grader transferred before the test was administered.

DOLORES DURKIN

Testing the hypothesis of no relationship between age, as defined by the three grade-levels, and kind of response, a chi square of 13.76 was obtained ($.005 < p < .010$). The data thus suggest that the different kinds of justice-concepts are significantly related to different CA levels.

As can be seen in Table 1, 27 subjects responded to Situation A in favor of a return of aggression. None in the group, however, approved of returning "a push besides" (Situation B). They unanimously denounced this behavior in such terms as: "It makes the fight unfair." "That other guy never pushed him." "He's asking for more."

TABLE I

KINDS OF RESPONSES AND NUMBER OF SUBJECTS GIVING THEM

Kind of Response	G R A D E		
	2	5	8
Tell authority person	15	13	27
Return identical aggression	8	15	4
Other	5	10	4
Ignore aggression		6	1
Withdraw from situation	3		2
Have aggressor apologize		2	
Tell aggressor to stop		2	
Exclude aggressor from play	1		
Do nothing			1
Undecided	1		

Six fifth graders and 12 eighth graders showed, through the questions they asked, concern for the particular circumstances of the situation being judged. They questioned the motive for the aggression ($N = 13$), the deliberativeness ($N = 3$), and the severity ($N = 2$) of the aggression. Testing the hypothesis of no relationship between age, as defined by grade-level, and this overt concern for particular circumstances, a chi square of 12.0 was obtained ($.001 < p < .005$). The data thus show a significant relationship between a subject's concern for particulars and his CA level.

Responses and IQ Level

Table 2 shows the number of subjects at two different IQ levels giving the various kinds of responses. It also lists the value of chi square and the level of significance obtained when hypotheses of no relationship between kind of response and IQ level were tested. On the basis of the chi square values, there appears to be no significant relationship between IQ level and kind of justice-concept in grades 2 and 8. A statistically significant rela-

CHILD DEVELOPMENT

TABLE 2

KINDS OF RESPONSES AND NUMBER OF SUBJECTS AT TWO DIFFERENT IQ LEVELS GIVING THEM

<i>IQ Level</i>	<i>Tell</i>	<i>Hit Back</i>	<i>Other</i>	<i>Chi Square</i>	<i>p</i>
<i>Grade 2 (N = 27)</i>				1.84	< .50
Above Median	5	5	3		
Below Median	9	3	2		
<i>Grade 5 (N = 38)</i>				7.42	< .05
Above Median	3	8	8		
Below Median	10	7	2		
<i>Grade 8 (N = 35)</i>				1.04	< .70
Above Median	14	1	2		
Below Median	13	3	2		
<i>Grades 2, 5, 8</i>				6.17	< .05
Above Median	21	16	13		
Below Median	33	11	6		

tionship does exist, however, in grade 5 and again when the three grade-levels are combined.

A chi square of 0.89 was obtained when the hypothesis of no relationship between IQ level and overt concern for particular circumstances was tested ($.25 < p < .50$). On the basis of this finding, the emergence of what Piaget calls "equity" does not appear to be significantly related to intelligence.

DISCUSSION

While subjects' responses in the present study do substantiate Piaget's contention of a relationship existing between chronological age and concepts of justice, they do not substantiate his more specific proposal that "children maintain with a conviction that grows with their years that it is strictly fair to give back the blows one has received" (9, p. 301). Only between grades 2 and 5 is there evidence of an increasing acceptance of reciprocity; eighth grade subjects, like those in grade 2, tend to seek justice in the authority person. This is not to say, of course, that similarity in grade 2 and grade 8 proposals reflects similarity in the two groups' disposition toward such a proposal. For the second graders, telling an authority person was a quick and apparently obvious solution to the problem at hand. For eighth graders, on the contrary, it was a solution that was proposed with neither haste nor enthusiasm. The kind of reasoning that led to their acceptance of it is graphically depicted in the verbatim response of one subject:

DOLORES DURKIN

You're not supposed to hit the guy, but you can't walk away from him or he'll always be doing that. You have to defend yourself, and if you tell the teacher they'll all be after you for telling. And if you hit him, the teacher will get mad at you. I don't rightly know exactly. I guess he should tell the teacher or report him in some way so it won't happen all the time.

Another kind of difference that was evident in the older subjects' responses was their overt concern for possible mitigating factors in the situation being judged. This finding substantiates Piaget's contention that older children's moral judgments are characterized by equity. It can be conjectured, of course, that many or even most of the subjects at all three grade-levels were actually concerned about the particulars of the situation they were judging; however, none of the second graders, six of the fifth graders (15.8 per cent), and 12 of the eighth graders (34.3 per cent) were sufficiently concerned to ask questions about them. For this total group of 18 subjects, hitting another child was not an aggressive act, and therefore not subject to punishment, if it were accidental, if it were lacking in severity, or if it represented a return of aggression previously received. But these conditions not being present, the group unanimously responded in favor of some punishment. It is these latter responses that are included in the summaries in Tables 1 and 2.

Findings concerning the function of intelligence in moral-judgment development suggest no clear-cut conclusions. When IQ scores are divided into above-the-median and below-the-median levels, the data show no significant relationship between intelligence and kind of justice-concept held at the grade 2 and grade 8 levels, but do suggest a relationship at the grade 5 level, and again when all three grade levels are combined.³ Even where the relationship is significant the trends are not consistent. At the grade 5 level, for example, there is a tendency for the less intelligent subject to seek justice in an authority person, and for the more intelligent subject to avoid both "telling" and hitting back. When the three grade levels are combined, however, it is the more intelligent subject who tends to solve the problem by telling an authority person while the less intelligent subject tends to favor either reciprocity or some other solution.

When IQ scores of the subjects showing concern for details were examined, again no significant relationship between intelligence and such concern was evident. Eleven in the group had IQ scores above the median; seven, below the median.

On the basis of these various findings it would seem that no definite conclusions can be drawn concerning the function of intelligence in the development of moral judgment until other studies involving more subjects

³ When IQ scores are divided on the basis of quartiles there appears, both within and between the groups, no significant relationship between the variables. This finding, though consistent throughout the grades, was not used as a basis for drawing conclusions since it was established, prior to the statistical analysis, that such a division would result in too small theoretical frequencies in the chi square tests.

CHILD DEVELOPMENT

are carried out. Other studies in this general area might also provide evidence to substantiate two incidental but nonetheless consistent findings in the present study, namely:

1. Children tend to be much more certain about what they should not do than they are about what they should do. Perhaps this finding reflects an emphasis on the negative side of moral codes to such an extent that the path to goodness is left highly undefined for the child.
2. Children tend to be greatly perplexed about what constitutes rightful self-defense and what constitutes unlawful aggression. Perhaps this finding only reflects the highly complicated and often conflicting attitudes toward physical aggression preached and practiced by adults.

SUMMARY AND CONCLUSIONS

Middle-class subjects of three different age-groups were questioned about the problem of restoring right order in instances of physical aggression between children. Their responses were examined in order to identify possible developmental trends in kinds of solutions proposed and, further, to compare such trends with those suggested by Piaget as being basic to the evolution of a sense of justice in the child. The function of intelligence in moral-judgment development was also examined.

Finding in the present study show that:

1. Piaget's contention of a relationship existing between chronological age and justice concepts is substantiated. However, the data do not support his more specific proposal that acceptance of reciprocity as a justice-principle increases with age.
2. In no instance did acceptance of reciprocity include approval of returning aggression that was different from the aggression received. This unanimous reaction duplicates the Piaget finding that children who approve of reciprocity do not accept "a sort of arbitrary punishment whose content bears no relationship to the punishable act."
3. Older children tend to show concern for possible mitigating factors in the situation being judged. This bears out Piaget's finding concerning the emergence of "equity" with increase in age.
4. The role of intelligence in moral-judgment development remains undefined. Data concerning the relationship of intelligence and kind of justice concept are conflicting. However, findings do support the hypothesis of no relationship between intelligence and "the feeling of equity."

REFERENCES

1. ABEL, THEODORA M. Moral judgments among subnormals. *J. abnorm. soc. Psychol.*, 1941, 36, 378-392.
2. BRUCE, M. Animism vs. evolution of the concept "alive." *J. Psychol.*, 1941, 12, 81-90.

DOLORES DURKIN

3. DEUTSCHE, JEAN M. The development of children's concepts of causal relations. *Univer. Minn. Inst. Child Welf. Monogr.*, 1937, No. 13.
4. HARROWER, M. R. Social status and the moral development of the child. *Brit. J. educ. Psychol.*, 1934, 4, 75-95.
5. LERNER, E. *Constraint areas and the moral judgment of children*. Menasha, Wis.: George Banta, 1937.
6. MCCARTHY, DOROTHEA. Language development in children. In L. Carmichael (Ed.), *Manual of child psychology*. New York: Wiley, 1946. Pp. 492-630.
7. MACRAE, D. The development of moral judgment in children. Unpublished doctoral dissertation, Harvard Univer., 1950.
8. OAKES, M. E. Children's explanations of natural phenomena. *Teach. Coll. Contr. Educ.*, 1947, No. 926.
9. PIAGET, J. *The moral judgment of the child*. London: Harcourt, Brace, 1932.

[The text on this page is extremely faint and illegible. It appears to be a multi-paragraph document, possibly a letter or a report, with several lines of text visible across the page.]

to
o
o
ti
ch
a
re
L
to
n
a
fo
s
g
ti
e
h
y
9
t
s
t
t
v
-
f
D
a
C

LEARNING BY IMITATION IN KINDERGARTEN CHILDREN

JUDY F. ROSENBLITH*¹

Harvard University,

This study asks four main questions: (a) Does having a model lead to significantly greater improvement in learning than additional experience only? (b) Is the extent of a child's learning by copying an adult "leader" or model affected by the sex of that leader (considered, especially, in relation to the child's sex)? (c) Does the way in which the leader treats the child immediately before the copying session affect learning? (d) Is there an interaction between the "sex" and "treatment" variables?

The problem of imitation examined in an experimental framework has received its most extensive treatment in Miller and Dollard's book *Social Learning and Imitation* (4). They demonstrated that the tendency to imitate is acquired in rats and children, and that the operations of reinforcement are necessary for the establishment of imitative tendencies.

Schein (8) attempted to test the assertions of Miller and Dollard with adult Ss. Studying the effect of reward on adult imitative behavior, he found that a significant number of Ss "learned" to imitate a model when such imitation was rewarded. He also found that the imitative response generalized to a similar but new situation. Since the absolute level of imitation was not high, Schein speculated about motives (e.g., "imitation is equivalent to cheating") which might have been operative to demand that his adult Ss not imitate. Such motives all appeared less apt to be present in young children than in adults. Also, Schwartz (9) found more imitation in 9- and 10-year-olds than in 15- and 16-year-olds, suggesting that more imitation might be expected in kindergarten Ss.

Miller and Dollard classified imitation into three categories of behavior: *same* behavior, *matched-dependent* behavior, and *copying*. They asserted that all three behaviors are learned via the principles of instrumental conditioning. Learning by imitation is linked to their concept of copying in which one person learns to model his behavior on that of another. They

* Department of Psychology, Brown University, Providence 12, Rhode Island.

¹ This paper is based on a dissertation submitted to Harvard University in partial fulfillment of the requirements for the Ph.D. The author would like to express gratitude to Dr. Wesley Allinsmith for his encouragement and counsel. This research was aided by a grant to the Laboratory of Human Development from the Public Health Service.

CHILD DEVELOPMENT

further asserted that all three kinds of imitative learning respond to the same conditions.

In this study there was no attempt to teach the child to copy. There was reliance on the child's having learned to make imitative responses as a result of past experience in the processes of socialization. Generalization of this established learning to imitate is what was looked for. As Miller and Dollard have said: "An analysis in terms of learning principles indicated that the factor of generalization should play an important role in the processes determining the degree to which any leader will be reacted to as a prestigious model" (4, p. 166). Thus, generalization to a male model from significant males in the child's life was expected, and similarly for a female model. Since the child has very little contact with the particular adults who serve as models, it was assumed that the degree to which the model is effective does stem from such generalization. Nevertheless, rewards were not excluded from the situation. The leader to be imitated or copied was not only prestigious by virtue of age and size, but the leader's skill was emphasized by verbally rewarding their correct performance. The child was also verbally rewarded when his performance was a sufficiently close copy of the leader's (i.e., was correct).

This problem can also be examined in the theoretical framework of identification theory. Doing so poses the question of whether an adult leader of the same sex as the child might not be a more effective model for imitative behavior for 5-year-old children.

Miller and Dollard, in commenting that sex-typing may belong to a list of social conditions that may enhance or inhibit imitation or the tendency to imitate, do not tell at which age level this factor should become important. It is, however, not unreasonable to assume that at kindergarten age most children have had more reinforcement for imitating a female and hence that a female model should be more effective for both sexes. Such a view would be congruent with that expressed by Miller and Dollard or by contiguity theorists with regard to the secondary reward value which people acquire.

Most identification theories, on the other hand, would lead one to expect that a good deal of identification has taken place by the kindergarten age. Thus, one would expect that a leader of the same sex as the child should be the most effective, with considerable individual fluctuation due to differences in the stage of identification that has been reached.

Another aspect of identification theories leads to a consideration of the role of threat in the formation of identification (Freud, castration threat; Fromm, threat from authoritarian father). Mowrer (5) posits two kinds of identification. The first is developmental, it is with the mother for both sexes, threat does not play a role in it, and it is related to the learning of skills. The second kind of identification is a later form. It is with the parent of the same sex and threat may play a role. Mowrer calls this kind defensive identification, and says that it is related to character learning and is well under way by kindergarten age. In line with these considerations one might

JUDY F. ROSENBLITH

expect differences in the amount of identification shown or generalized (hence in the effectiveness of the adult as a model for copying) according to the presence or absence of an implied threat in the adult-child relation. The contingency or uncertainty-reduction theory of effectiveness of secondary rewards would lead to the same set of expectations.

These considerations led to the following experimental manipulations: (a) the adult leader was attentive to the child for the entire period prior to the copying session for some Ss, and (b) the adult withdrew attention from the child after half of the period and stayed withdrawn for the remainder of the period prior to the copying session for other Ss. Such withdrawal presumably constitutes a mild threat to the child's relation to the adult.

The present study was not intended to provide an *experimentum crucis* for deciding the merits of the above theories, but only to shed light on these problems.

PROCEDURE

The Ss were 120 kindergarten children from two public schools in a Boston suburb composed of middle and upper-middle class residents. Each child was brought individually from the classroom to the experimental room by the E in order to "play a game." In this initial session the child was tested using the Porteus Maze Test (6); this test was administered with slight modifications in the instructions and procedure to make it more suitable to the age of the children.² The Porteus Test contains one maze for each year from 3 through 11. Two trials were allowed on each of these as in the Porteus Test. There are also mazes for 12- and 14-year-olds and adults. Porteus allows four trials for each of these mazes. In this experiment only two trials were allowed since the 5-year-old Ss who reached the 12-year-old level or above were apt to be frustrated by four trials on such difficult mazes and want to quit the "game." The Porteus criteria for discontinuing testing were used, i.e., failure on two consecutive or on two out of three consecutive age levels.

On the basis of level of performance in this initial session children were assigned to one of four groups. Group 1 consisted of those who passed only the first two mazes, (i.e., the 3- and 4-year mazes). Group 2 consisted of those who passed a total of three or four of the mazes. Group 3 passed five mazes, and Group 4 passed six or more of the mazes. Children from each group were then assigned to each of the treatments.

One to three weeks after this initial session the child was brought from his room by E to "have a second turn at the game." This will be referred to as the copying session. Just before arriving at the experimental room, E said to each child who was not in a control group: "You know what? I have someone else here to play the game with us." E and child then entered the room and E introduced the child to the other adult by first

² Copies of the exact procedure may be seen in (7).

CHILD DEVELOPMENT

names. *E* then explained to the child that she had something to do for a while and invited the child and the other adult to go to another table in the room and play with the toys on it until *E* was ready to play the game.³ The child was seated facing the table of toys and the adult at the adjacent side facing the child and/or table.

It should be noted that *E* essentially was treating the other adult in the same manner as the child. *E* did nothing to enhance the adultness or maleness (or femaleness) of the leaders except to reward their proficiency. They were not called Mr. or Mrs. and not described as fathers or mothers. However, they were adult and they were proficient at the task. During the play period the leader behaved as an adult interested in the play of the child and not as another child playing.

The experiment was first run using a male for a leader. The treatments or conditions were: (I) Male leader with a child of the same sex, leader attentive to child throughout the 10-minute period that preceded the copying session proper. (II) Male leader with a child of the same sex, leader attentive to the child for 5 minutes, then leader said: "I'm sorry, I can't play with you any more, I have to read a book." Leader then turned his chair so his back was toward the child and read a book for 5 minutes. (III) Control Ss, males, who were brought to the room by *E* and who immediately received two additional trials at each maze previously failed, until the criterion for discontinuing testing was reached. (IV) Male leader with a child of the opposite sex, attentive throughout (as in I). (V) Male leader with a child of the opposite sex, withdrew attention after 5 minutes (as in II). (VI) Control Ss, females, tested in the same fashion (as in III). The entire design was then repeated using a female leader.

In treatments I, II, III, IV, and V, at the end of the 10-minute period *E* said: "I'm ready to play the game now." Adult leader and child moved to the *E*'s table. They took adjacent seats at the table on the side opposite *E* and facing her with the adult to the right of the child. *E* then said to the child: "You've played this game with me before. — (adult) wasn't here when we played then, so how would it be if you take turns and give — (adult) the first turn?" (We might note here that no child ever argued with this suggestion of taking turns.) The adult was then given the first maze which the child failed previously and the same instructions that the child had received. The leader did the maze correctly at a rather slow pace and with pauses at choice points accompanied by some visual search. The child watched during this period and then was given a turn at the same maze. To ensure that the child was watching the entire performance, the adult paused when he was not and only continued when the child had again turned his attention to the maze. There were two such

³ The toys on the table were chosen for their masculine and feminine identification potential. *E* was actually recording the child's behavior as the child played with the toys. Ratings of these protocols for masculine and feminine identification will be examined later. The toys included shaving set and make-up kit, carpenter tools and kitchen utensils, and a family of dolls, among others.

JUDY F. ROSENBLITH

trials on each maze failed initially (unless the child passed it correctly on the first trial, in which case there was only one), until the criterion for discontinuing testing was reached. This yields a design in which there are six treatments (including controls) and four replications (by initial level of performance). The part of the study using the male leader was started at one school and completed at the second. The repetition with the female leader was started at the second school and completed at the first.

The number of Ss decided on for each cell was based on the distribution of passes obtained by the children in the first school. Since the second school had a different distribution of passes (fewer children with a high number of passes), the number of children from group 4 of initial performance level in each cell was decreased from 3 to 2, thus dropping the total *N* from an anticipated 132 to 120 (see Table 1).

TABLE 1
EXPERIMENTAL DESIGN
(Number of Ss in each cell is shown)

Initial Performance Groups	MALE Ss			FEMALE Ss			
	A†(I)	WA†(II)	C†(III)	A(IV)	WA(V)	C(VI)	
	M A L E L E A D E R *						
1	3	3	3	3	3	3	18
2	3	3	3	3	3	3	18
3	2	2	2	2	2	2	12
4	2	2	2	2	2	2	12
	10	10	10	10	10	10	60

* The entire design is repeated with a female leader.

† A stands for leader attentive throughout, WA, for withdraws attention for last half of period, and C, for controls in this and all subsequent tables.

The data to be analyzed here, in investigating the amount of learning by imitation, are the number of new mazes passed ("new passes") in the second or copying session.

RESULTS

Since the distribution of the number of new passes is not normal, and since the number of mazes passed forms only an ordinal scale, the results will be examined using nonparametric statistics.

The data in Table 2 refer to the over-all experimental design, i.e., to both the male and female leader experiments. If one applies Friedman's χ^2 , (9) to these data, one finds the *p* values⁴ shown in Table 3. The differences found justify separate examination of pairs of treatments.

⁴ All *p* values reported in this paper are for a two-tailed test.

CHILD DEVELOPMENT

TABLE 2

NUMBER OF NEW PASSES OBTAINED BY Ss IN THE VARIOUS CELLS

Initial Performance Groups	MALE Ss			FEMALE Ss		
	A (I)	WA (II)	C (III)	A (IV)	WA (V)	C (VI)
M A L E L E A D E R						
1	7	6	5	5	4	2
2	9	12	8	9	6	6
3	4	6	1	4	4	2
4	6	7	1	4	5	2
Total	26	31	15	22	19	12
F E M A L E L E A D E R						
1	6	2	4	4	1	1
2	10	4	6	6	3	4
3	8	4	2	3	3	4
4	6	4	1	3	5	3
Total	30	14	13	16	12	12
Total	56	45	28	38	31	24

Before doing so, it is interesting to examine the data according to the sex of Ss and of leaders. Table 3 shows the p values of χ^2_r calculated for the various combinations of sex of Ss and of leader. Male leader and male Ss show more influence of these manipulations.

Examination of pairs of treatments by a method described by Wilcoxon (11) shows that boys and girls who are control Ss do not differ significantly, although there is a tendency for the control boys to improve more than the control girls. This tendency is close to the .05 level for those whose

TABLE 3

COMPARISON OF TREATMENTS EXAMINED BY THE FRIEDMAN χ^2_r

Treatments Compared	<i>p</i> V A L U E S	
	Male Leader	Female Leader
Total Table (Conditions I through VI)01	.09
Treatments—controls not included (Conditions I, II, IV, V)06	.052
Conditions involving male Ss (Conditions I, II and III)04	.05
Conditions involving female Ss (Conditions IV, V and VI)15	.80

JUDY F. ROSENBLITH

initial performance placed them in level 1. The remaining results of these comparisons are presented in Table 4. In summary it can be seen that: (a) having a model is more effective than merely having additional trials; (b) girls seem less sensitive to the experimental manipulations than are boys; (c) the female leader was less effective than the male leader; (d) attention is more effective than withdrawal of attention.

TABLE 4
COMPARISONS OF PAIRS OF TREATMENTS USING THE
WILCOXON RANKING METHOD

Comparison†	p VALUES*			
	MALE LEADER		FEMALE LEADER	
	Male Ss	Female Ss	Male Ss	Female Ss
A plus WA vs C	.01 (A&WA)	.01 (A&WA)	ns (A&WA)	ns (A&WA)
A vs C10 (A)	.05 > p > .02 (A)	.10 (A)	ns (A)
WA vs C05 (WA)	.10 > p > .05 (WA)	ns (WA)	ns (WA)
A vs WA	ns (WA)	ns (A)	ns (A)	ns (A)
A plus WA vs C (by initial performance groups)				
Group 1	ns (A&WA)	.05 (A&WA)	ns (C)	ns (C)
Group 205 (A&WA)	ns (A&WA)	ns (C)	ns (A&WA)
Group 3 and 4	.05 > p > .02 (A&WA)	ns (A&WA)	.01 (A&WA)	ns (A&WA)

* Exact p values or p values larger than the .10 level are not available in the published tables. "ns" is used to indicate that p is greater than .10, i.e., is not significant.

† The abbreviations in the parentheses indicate which side of the comparison gave the superior performance.

Certain qualifications of these general statements are found in the detailed results. The finding concerning the effectiveness of the male and female leaders cannot be labeled a sex difference without further replication since it is possible that individual characteristics of the leaders rather than their immediately perceptible sex characteristics were responsible. In any event, these rather major effects of sex (of both leader and child) which operate in a way that cuts across the original experimental design make the next step of analysis less apt to yield significant results.

The next step was to examine the two experiments together. Since the Ss were not randomly assigned to treatments across both sexes of leader, some justification for combining the data is probably needed. Both leaders did have Ss from both schools, though not equal proportions of Ss from each. The control Ss in both replications showed no significant difference; in fact, the rank totals are 412 for the first group and 408 for the second. It thus seems reasonable to conclude that the two populations are similar enough to justify combining them for comparisons.

CHILD DEVELOPMENT

Analyzing the data in this way, the design yields five conditions: opposite sex attentive, opposite sex withdraws attention, same sex attentive, same sex withdraws attention, and controls. If these treatments are examined over the four replications provided by initial level groups, the χ^2_r is 8.85, which (with 4 *df*) has a *p* value between .06 and .07. As in the earlier analysis, the boys show a significant effect of treatments (*p* is .02) and the girls do not (*p* is .15). An extension of the Kruskal-Wallis technique⁵ yields an analysis of variance the results of which are shown in Table 5.

TABLE 5
KRUSKAL-WALLIS ANALYSIS OF VARIANCE AS EXPANDED BY HYMAN

Source of Variation	<i>H</i>	<i>df</i>	Approximate <i>p</i>
Among subcells	36.6	39	<i>ns</i> (.5)
Treatments	7.87	4	.10
Levels of initial performance	7.13	3	.07
Sex of Ss	3.00	1	.09
Sex \times Levels72	3	<i>ns</i> (.85)
Levels \times Treatments	8.4	12	<i>ns</i> (.75)
Sex \times Treatments	7.43	4	.11
Sex \times Treatments \times Levels	2.05	12	<i>ns</i> (.99)

DISCUSSION

The difference between the results for the male leader and the female leader are striking, as are the differences between the results for boys and girls. The fact that the boys tended to show more improvement than the girls may provide more ceiling for finding differences between treatments for the boys. However, such differences between boys and girls remain puzzling since they had been equated for pretest performance. Although there was no significant difference between boys and girls as controls, the consistent superiority of all Ss with a male leader is noted ($p < .05$ that they differ from controls by chance, but $p < .10$ that Ss with the female leader do). Hartup (3) found that boys learned a concept formation task more quickly than girls. He speculated that the blocks involved were more salient for boys than for girls. (Note: The mazes used as a subtest on WISC correlate most highly with block design and object assembly, about .48). Publications on the Porteus Test fail to give any information on sex differences that sheds any light on this problem.

⁵ This extension of the Kruskal-Wallis analysis of variance was suggested to the author by Dr. R. Hyman of the Department of Social Relations at Harvard.

JUDY F. ROSENBLITH

One should also note that the level of performance on the pretest is a significant source of variation. Thus, it is wise to control for level of pretest performance in a study such as the present one. The fact that levels did not enter into any significant interaction term mitigates this conclusion. It does, however, seem (*see* Table 4) more profitable to look for certain effects of treatments at particular levels rather than in a random population or a representative one.

Now one can turn to the questions raised at the beginning of the paper.

1. Does having a model have a significant effect on the amount of improvement? The answer to this question is yes.

2. Is one sex of leader more effective in evoking learning by imitation? In particular the question raised was whether a female leader or a leader of the same sex as the child would be more effective. Actually the male leader was more effective for both sexes, though only significantly so for the girls. One possible explanation for this finding might lie in the fact that an adult male is more unusual in the school setting, hence has more reward value. Gewirtz and Baer (2), in studying the effectiveness of a social reinforcer (approval) in conjunction with social deprivation, found that the effectiveness of the isolation was qualified by an interaction which indicated that the woman was less effective relative to the man under the condition where *Ss* were tested immediately on removal from the play group than the condition where *Ss* were tested after a 20-minute period of total isolation. Our situation would seem to be closer to the former condition. Perhaps the whole school setting is one of social deprivation in respect to adult males who play a role in the life of kindergarten children. In summary, there is an effect of sex of leader, but it is not the effect anticipated from either of our theoretical approaches.

3. Does an implied threat to the relation between the child and the adult operate to enhance identification and thus enhance learning by imitation? For boys there was a tendency for those from whom the male leader withdrew attention to do better. They were not significantly better than boys with an attentive male leader; but they were better than controls at the .05 level, while the male leader attentive group were better than controls at only the .10 level. For girls, however, the effectiveness of the leader of the same sex withdrawing attention does not appear. With the female leader, as with the male, girls tend to do better when they have attention throughout.

The failure to find a relation between implied threat and better learning for girls with a female leader is particularly puzzling in view of Hartup's findings (3). In his study he found that 4-year-old girls learned a simple concept formation task twice as fast when the woman for whom they were doing it had been very nurturant for 5 minutes and then had sat "busy" at her desk for 5 minutes. This quicker learning was not shown for the 4-year-old boys in general. There are a number of differences that could account for the discrepancy: (a) He used 4- instead of 5-year-olds. (b) His

CHILD DEVELOPMENT

task was presented by the person who withdrew instead of by a third person. (c) There were differences in the task (his Ss learned the concept that determined which of four blocks is the "right" one, ours the way to do a maze). (d) Our results also might be idiosyncratic to our female leader.⁶ Hartup used two females and found no significant differences between them. (e) The nurturance which his leaders interrupted was of a more active kind (getting down on the floor and playing with the child). The possible effects of the last four differences are very difficult to speculate about. There is, however, an interesting question we might raise concerning the first point. There are reasons to expect on the basis of Freudian thinking that a girl changes her primary object from mother to father at about age 4. If that is the case, the girl may be particularly sensitive to threat from females at that age. Boys presumably do not change until resolution of the Oedipus complex, well into the fifth year on the average. In that case, 5-year-old boys might be particularly sensitive to threat from males. This interpretation would be congruent with both Hartup's and our findings. Very general support for this idea may be found by examining the data for all girls and boys in the sample who are under 5 years, 3 months, in comparison with all those who are over 5 years, 7 months. We find that the older boys improve significantly more than the younger ($p = .02$), while the older girls do not improve significantly more than the younger ($p > .10$). Unfortunately, there are too few cases to enable us to look at the operation of the attention and withdrawal of attention variables within these age subgroups.

In summary, there is equivocal evidence for threat having the effect of enhancing performance where boys with the male leader are concerned, but in no other case.

4. Is there an interaction between the sex and treatment variables? Yes. This answer must be qualified by saying that we cannot be certain whether it is the sex of the leaders or other personal characteristics which account for some of the interactions. Further replications are needed to settle this issue. However, one should bear in mind that generalization from adults significant in the child's life was expected to be more important than the personal characteristics of the leaders to whom the Ss were exposed for only a brief span of time. Also, Hartup's failure to find differences between two female adults in their effect on the children, and the lack of differences in our own small check are signs which indicate that this is a reasonable view.⁷

⁶ We did a spot check using another female leader for 4 children (2 boys and 2 girls, all at the same initial level of performance). It is suggestive that there was no apparent difference between results for these 4 children and a comparable 4 from our female leader experiment.

⁷ Additional evidence of this kind is provided by Bishop (1), who found that when children were brought into a play room situation with a neutral adult (female) they reacted in the same way they had earlier reacted to their mothers with respect to aggressive stimuli, cooperation, noncooperation and resistance.

JUDY F. ROSENBLITH

On the basis of the data, boys and girls both respond better to attention from the opposite sexed leader than to withdrawal. Boys, and not girls, tend to do better after withdrawal of attention from the leader of the same sex. As indicated above, the differential functioning of threat may be, in part at least, a function of the age of the Ss. The rank orders of effectiveness of treatments for the over-all design is given in Table 6.

TABLE 6
RANK ORDERS OF AMOUNT OF IMPROVEMENT

	<i>SS-A</i>	<i>SS-WA</i>	<i>OS-A</i>	<i>OS-WA</i>	<i>C</i>
Boys	3 (26)	1 (31)	2 (30)	5 (14)	4 (15)
Girls	3 (17)	4 (12)	1 (21)	2 (19)	5 (11)

NOTE. — *SS-A* stands for leader of same sex—attentive throughout;
SS-WA stands for leader of same sex—withdraws attention;
OS-A stands for leader of opposite sex—attentive throughout;
OS-WA stands for leader of opposite sex—withdraws attention;
C stands for controls, additional experience with no leader.
 Figures in parentheses represent the number of new passes obtained by all subjects in that cell.

SUMMARY

The effectiveness of learning by imitation was studied in a context which permitted examination of a number of variables relevant to learning and identification theories. These were: (a) the effectiveness of having a leader or model as contrasted with experience in the absence of a model; (b) the effectiveness of the sex of the leader and of the leader's sex in relation to that of the child; (c) the effectiveness of the adult leader who gives attention to the child for the entire period preceding the imitation, as contrasted with the adult who pays attention to the child for half the time and withdraws attention for the remaining half.

In general, having a model was more effective than merely having additional trials. There were important differences between the effectiveness of the male leader and the female leader. The male leader was, in general, more effective. There were also important differences between boys and girls. Boys showed more improvement. Girls seemed less sensitive to the experimental manipulations. There was a tendency for attention to be more effective than withdrawal of attention except in the case of boys with a male leader. The specific findings were examined in detail and their relation to current theories discussed. Analysis of variance on the male and female leader parts of the study combined showed effects of: (a) treatments, (b) sex of Ss, (c) initial or pretest level of Ss' performance on the mazes, and (d) interaction between the sex of Ss and the treatments.

CHILD DEVELOPMENT

REFERENCES

1. BISHOP, BARBARA M. Mother-child interaction and the social behavior of children. *Psychol. Monogr.*, 1951, 65, No. 328.
2. GEWIRTZ, J. L., & BAER, D. M. Does brief social "deprivation" enhance the effectiveness of a social reinforcer ("approval")? *Amer. Psychol.*, 1956, 11, 428-429.
3. HARTUP, W. W. Nurturance and nurturance-withdrawal in relation to the dependency behavior of preschool children. *Child Development*, 1958, 29, 191-201.
4. MILLER, N. E., & DOLLARD, J. *Social learning and imitation*. New Haven: Yale Univer. Press, 1941.
5. MOWLER, O. H. *Learning theory and personality dynamics*. New York: Ronald, 1950.
6. PORTEUS, S. D. *The Porteus Maze Test and intelligence*. Palo Alto: Pacific Books, 1950.
7. ROSENBLITH, JUDY F. Imitation in kindergarten children. Unpublished doctoral dissertation, Radcliffe College, 1958.
8. SCHEIN, E. H. The effect of reward on adult imitative behavior. *J. abnorm. soc. Psychol.*, 1954, 49, 389-395.
9. SCHWARTZ, N. An experimental study of imitation. The effects of reward and age. Senior honors thesis, Radcliffe College, 1953.
10. SIEGEL, S. *Nonparametric statistics for the behavioral sciences*. New York: McGraw-Hill, 1956.
11. WILCOXON, F. Individual comparisons by ranking methods. *Biometrics Bull.*, 1945, 1, 80-83.

AUDITORY LOCALIZATION AND ADAPTATION TO BODY TILT: A DEVELOPMENTAL STUDY

ROBERT S. LIEBERT* and RITA G. RUDEL*

New York University College of Medicine

This study represents an effort to demonstrate that certain developmental changes which have been of considerable interest in vision have their parallel in audition. The comparison of performances in each modality may help to delineate the relative contributions of central and peripheral factors in perception.

Ninety-six years ago, Aubert (1) found that an adult, in the dark, could adjust a luminescent line to a position of verticality with considerable accuracy as long as his head and body were erect. However, when *S* tilted his head, constant errors appeared in the localization of the line, and these persisted as long as the room was dark. These effects of body tilt on perception have been studied intensively by Werner and Wapner (10). They found that when *S* is tilted in a chair about 30° to either side, the line will appear vertical to him when it is in fact displaced by several degrees in a direction opposite his body tilt. Thus, when the observer is tilted to the left, the line appears vertical when it is several degrees to the right of vertical. This paradigmatic situation of visual judgments of verticality has been of particular value in dealing with problems of perceptual development and pathological perception (2, 3, 6, 8). More recently, comparable effects were demonstrated in blindfolded *Ss* who adjusted a rod to a position of apparent verticality by palpating it while tilted (7).

This parallelism of effects in visual and "haptic" situations suggests but does not prove that the important factors are central rather than peripheral. Nevertheless, the visual effects could be somehow produced by

* Psychophysiological Laboratory, New York University—Bellevue Medical Center, 550 First Avenue, New York 16.

CHILD DEVELOPMENT

ocular rotation during body tilt, and the haptic effects by postural reflexes influencing the muscular tone of the arms and hands.

Such relatively peripheral factors are made considerably less likely by an extension of these studies to yet another sensory modality: Teuber and Liebert (5) developed an auditory analogue to these visual and haptic situations. They found that the normal adult is able to localize an ambient sound, moving in an arc in his coronal plane, with remarkable accuracy at the midline of his head, provided he be seated upright. When tilted to the left or right, however, he tends to displace the sound consistently in a direction opposite his body tilt. Thus, tilted 28° to his left, *S* reports the sound as being directly in his median plane, when it is actually more than 4° to the right of the median plane, and, conversely, when tilted to the right, the sound appears to him in the median plane when it is 5° to the left of it.

It should be noted, however, that this experiment with ambient sound differs in one respect from the visual or haptic settings of the vertical. The localization is referred to *S*'s own head, a point which moves as *S* moves, in contrast to the objective standard of verticality in space, independent of *S*'s position, in the visual and haptic experiments. In spite of these differences, Teuber and Liebert conclude that some central rather than peripheral mechanisms are involved. Otherwise, it would be hard to understand why there should be essentially analogous displacements during tilt in visual, haptic, and auditory localization. This interpretation could be strengthened further if one could demonstrate that auditory settings change with age in the same manner in which visual settings have been shown to vary systematically from child to adult.

Wapner and Werner (8), testing 237 children aged 6 to 19 for the effects of body tilt on apparent verticality, found the following: (a) There is a systematic change in the effect of body tilt on apparent vertical, depending on the age level; with development, the apparent vertical shifts from a position close to true vertical, but to the same side as body tilt, to a position distinctly opposite body tilt. (b) There are developmental changes in the effect of so-called "starting position." By "starting position effect" is meant the influence exerted by the position of the rod at the beginning of the trial (starting position) on the final position in which *S* perceives it as vertical. The apparent vertical is displaced from the plumb line in the direction toward the position at which the rod is initially placed. This starting position effect is greatest for the youngest children and decreases with age.

The findings of that study were interpreted within the framework of Werner's developmental theory (9), which postulates a progressive articulation between self (body) and object. Thus, at early age levels the position of the apparent vertical should be egocentrically tied to the body posture; the developmental shifts of the apparent vertical (to positions of increasing displacement to the side opposite tilt) reflect a decrease of such egocentricity, or relative independence of external space from changes of body

position. Similarly, Piaget (4) has suggested that the increase in size constancy with age (up to the overconstancy often found in the adult) is a function of an increasing differentiation between self and surround.

Interpretations such as Werner's or Piaget's are obviously based on the assumption that central rather than peripheral mechanisms underlie the contrasting effects in the visual situation, and that the same general features should be found in other modalities. Experiment I was designed to test this prediction.

EXPERIMENT I

Subjects and Method

Subjects were 36 boys and 36 girls, ranging in age from 5 years, 7 months to 17 years, 6 months. Most of them were permanent residents of Cape Cod. Ss were first instructed, then blindfolded and seated in a tilting chair to which the sound source was attached, a single PDR-10 phone emitting 10 millisecc. clicks generated once per sec. The phone could be moved in an arc in S's coronal plane, at a constant distance of 12 inches from the occipital pole. The clicks were presented in balanced order in four starting positions: 10° and 30° to the left, and 10° and 30° to the right of S's midline. Ss had to direct E to adjust the position of the source until it sounded to him as being directly above the midline of his head. This was done in three body positions: upright, 28° to the left, and 28° to the right. The order of these body tilts was also balanced to permit separate assessment of effects of tilts, and of starting position of source, on the midline localization of the sound. For purposes of analysis the groups of Ss of each sex were divided into six subgroups, each covering a 2-year span: 5 years, 7 months to 7 years, 6 months; 7 years, 7 months to 9 years, 6 months; . . . 15 years, 7 months to 17 years, 6 months.

Results

Body tilt. The *F* tests are presented in Table 1. Table 2 gives the mean position of the midline estimates under the three conditions of body tilt for each age group and for either sex.

The pattern of the results was very similar to that reported for the visual setting of a rod to vertical by Wapner and Werner's children. When judgments were made from the upright position, the youngest and oldest Ss showed, on the average, no significant deviation from the midline. With judgments made from a 28° -tilt to the left, there was a progressive displacement, with age, of the subjective auditory midline to 6.0° in the direction opposite tilt, that is, to the right. Conversely, there was a progressive displacement to 5.7° to the left with judgments made from a 28° -tilt to the right. The older the child, the greater was the constant error of localization; in other words, the young child tended to localize the auditory midline in much the same way, whether he was seated erect or tilted to one side, in contrast to the older groups. This progressive trend towards displacement is illustrated in Figure 1.

CHILD DEVELOPMENT

TABLE 1

SUMMARY OF ANALYSIS OF VARIANCE OF EFFECTS OF AGE, SEX, BODY TILT,
AND STARTING POSITION ON PERCEPTION OF AUDITORY MIDLINE

Source of Variation	df	SS	MS	F
Between Ss	71	1,728.16		
Betw. Age Groups (G)	5	79.15	15.83	< 1
Betw. Sex (S)	1	2.14	2.14	< 1
Interaction (G × S)	5	19.35	3.87	< 1
Error Between	60	1,627.52	27.13	
Within Ss	792	97,199.58		
Betw. Body Position (B)	2	5,977.26	2,988.63	14.05 ^{†**}
Betw. Starting Pos. (P)	3	61,407.96	20,469.32	105.25 ^{†***}
Interaction				
B × P	6	484.62	80.77	2.46 [§]
B × G	10	2,126.85	212.68	6.47 ^{§***}
B × S	2	29.96	29.96	< 1 [§]
P × G	15	2,917.21	194.48	5.92 ^{§***}
P × S	3	117.98	39.33	1.20 [§]
B × G × S	10	336.24	33.62	1.02 [§]
P × G × S	15	737.89	49.19	1.50 [§]
B × P × G	30	629.61	20.99	< 1 [§]
B × P × S	6	159.13	26.52	< 1 [§]
B × P × G × S	30	578.22	19.27	< 1 [§]
Within Error	660	21,696.65	32.87	
Total	863	98,927.74		

* $p < .05$.** $p < .01$.*** $p < .001$.

† Tested over MS of B × G.

‡ Tested over MS of P × G.

§ Tested over MS of Within Error.

TABLE 2

EFFECT OF BODY TILT ON MEAN POSITION OF PERCEPTION
OF AUDITORY MIDLINE*

Age	B O D Y P O S I T I O N		
	28° Left	Erect	28° Right
5-7	+2.12	-0.33	-0.77
7-9	+0.69	-0.56	-0.37
9-11	+3.14	+1.50	-2.10
11-13	+4.35	+1.10	-3.52
13-15	+5.01	+0.40	-4.85
15-17	+5.95	+0.64	-5.72

* A plus sign indicates a position of the estimate of midline to the right of the objective midline; a minus sign indicates a position to the left.

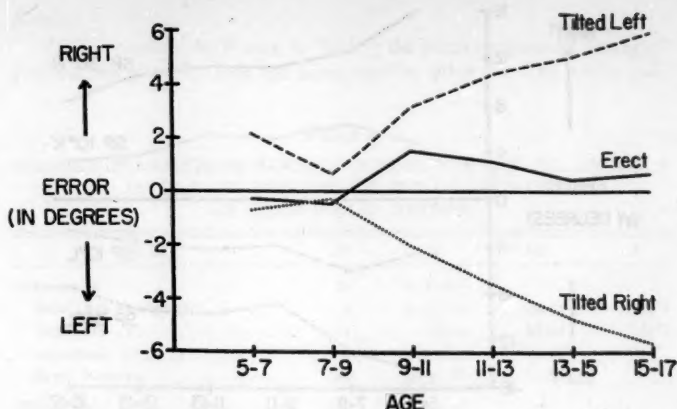


FIGURE 1—Effect of body tilt (left, erect, right) on localization of auditory midline as a function of age.

Starting position. The increasing displacement of the subjective auditory midline with age was accompanied by a decreasing dependence upon the starting position of the sound source in each trial. Table 3 shows the effect of starting position on the estimate of midline (body positions pooled) for each group. It reveals that the starting position error of the oldest group was just half that of the youngest group, when the sound source was presented 30° to the right or left. It might also be noted that all Ss, children and adults, approximated the objective midline much more accurately when the source originated 10° to either side than with starting positions of 30° .

TABLE 3
EFFECT OF STARTING POSITION ON MEAN POSITION OF PERCEPTION
OF AUDITORY MIDLINE*

Age	STARTING POSITION			
	30° Left	10° Left	10° Right	30° Right
5-7	-14.62	-4.70	+4.67	+16.01
7-9	-12.93	-6.20	+6.39	+12.40
9-11	-8.90	-3.98	+5.10	+11.12
11-13	-9.70	-4.39	+5.50	+11.17
13-15	-9.45	-3.84	+3.61	+10.42
15-17	-7.59	-2.61	+3.20	+8.17

* A plus sign indicates a position of the estimate of midline to the right of the objective midline; a minus sign indicates a position to the left.

CHILD DEVELOPMENT

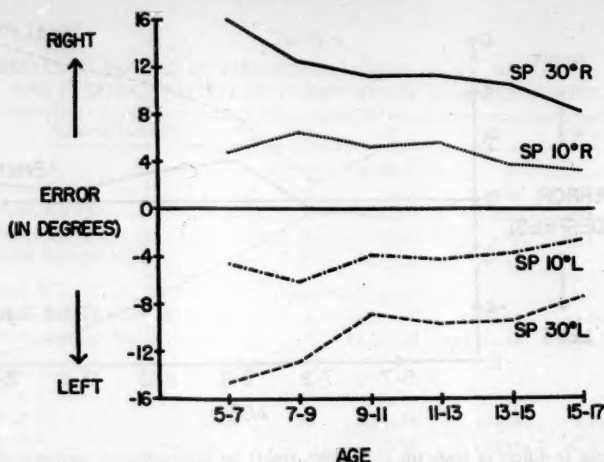


FIGURE 2—Effect of starting position of stimulus (30° and 10° left and right) on localization of auditory midline as a function of age.

to the right or left. Thus, when the sound originated at 30° to either side of the midline, the 5- to 7-year-olds were an average of 10.0° to the right or left of it in their judgments, while the 15- to 17-year-olds were only 5.4° away, thereby estimating the midpoint within a much smaller arc. Figure 2 graphically presents the decrease in starting position effects with increasing age.

Comment (Experiment I). Constant errors such as the one found for older Ss in Experiment I are often interpreted as compensatory or "counteractive." By placing the sound source beyond the objective midline, while Ss body is tilted to the other side of it, S commits an error through overcorrection for the angle of his tilt, a performance analogous to "overconstancy." The absence of this effect in the younger Ss could mean that these Ss are more tolerant of abnormal body posture, or adapt to it more readily than older children or adults. A second experiment was therefore undertaken to test whether increasing compensatory displacement of the auditory midline reflects decreasing adaptation to unusual body posture.

EXPERIMENT II

The second experiment was carried out after Experiment I, with the same 72 Ss. Each S, while blindfolded, was tilted six times; for 15, 30, and 45 secs., 28° to the left, and 28° to the right, in balanced order. He was slowly brought back to a position in which he said he felt perfectly upright. There was a 2-minute rest before each trial.

Results

Table 4 presents the *F* tests. In Table 5 the mean estimates of "upright" positions are given for each age group and for either sex. The results indi-

TABLE 4
SUMMARY OF ANALYSIS OF VARIANCE OF EFFECTS OF AGE, SEX, DIRECTION
OF TILT AND DURATION OF TILT ON ESTIMATE
OF ERECT SEATED POSITION

Source	df	SS	MS	F
Between Ss	71	8,187.07		
Betw. Age Groups (G)	5	3,547.60	709.52	11.71**
Betw. Sex (S)	1	88.02	88.02	1.45*
Interaction (G × S)	5	914.57	182.91	3.02†
Error Between	60	3,636.88	60.61	
Within Ss	360	3,536.50		
Direction of Tilt (T)	1	135.56	135.56	4.85‡
Duration (D)	2	49.46	24.73	2.58‡
Interaction				
T × D	2	3.08	1.54	< 1 ‡
T × G	5	55.87	11.17	< 1 ‡
T × S	1	1.45	1.45	< 1 ‡
D × G	10	82.93	8.29	< 1 ‡
D × S	2	1.42	0.71	< 1 ‡
T × G × S	5	139.83	27.97	2.92*
D × G × S	10	50.61	5.06	< 1
T × D × G	10	56.95	5.70	< 1
T × D × S	2	12.29	6.14	< 1
T × D × S × G	10	71.93	7.19	< 1
Error Within	300	2,875.12	9.58	
Total	431	11,723.57		

* $p < .05$.

** $p < .001$.

† Tested over MS of Error Between.

‡ Tested over MS of T × G × S.

‡ Tested over MS of Error Within.

TABLE 5
ERRORS IN ESTIMATE OF ERECT SEATED POSITION
(Mean Position in Degrees to Side of Tilt)

	A G E					
	5-7	7-9	9-11	11-13	13-15	15-17
Error	13.64	9.13	8.54	7.47	5.81	4.66

CHILD DEVELOPMENT

cate that there is a continuous decrease in the righting-self error with age. On the average, the youngest group, the 5- to 7-year-olds, were still 13.6° off the true vertical to the side of their original tilt when they reported feeling upright, whereas the oldest group, the 15- to 17-year-olds, were tilted only 4.7° when they felt upright. This difference was significant at the .001 level. There was no significant difference as a function of duration of time that *S* was kept in the tilted position (15, 30, and 45 secs.); nor were there any consistent differences that could be attributed to sex. The changes in self-righting with age are graphically illustrated in Figure 3.

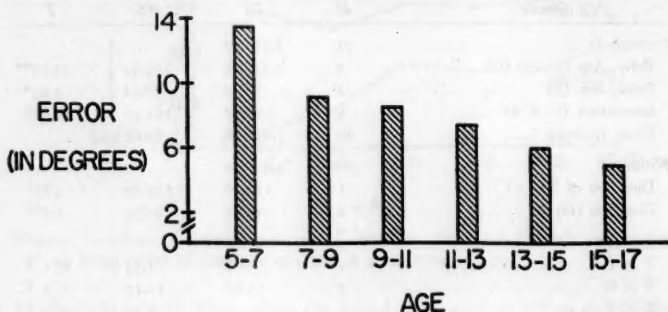


FIGURE 3—Mean error in estimate of erect seated position after body tilt as a function of age.

DISCUSSION

The findings, all of which are significantly related to age, as well as to each other, are: (a) increasing displacement of the estimated auditory midline towards the side opposite body tilt, (b) decrease in "starting position error," and (c) decreasing error of estimating the upright body position following brief suspension in a tilted position.

The increasing "counteractive error" that is shown by the older children in the estimate of auditory midline might well reflect a progressive decrease in adaptation with age to abnormal posture. The more tilted one feels, the more one compensates for the tilt—hence, the increasing "counteractive" error in older children, and the concomitant decrease of error in "righting-self" after body tilt.

The basis of the older children's "overcompensation" in estimating the midline, when tilted, might then be considered a form of overreactivity in a normally present compensatory mechanism. In order to keep our orientation in space constant despite variations in body position and movement, it is necessary to compensate for the concomitant changes in sensory stimulation. Without such compensation every displacement of the body would change the world about us. In the light, in everyday circumstances, this

mechanism works quite accurately, for it is governed by many other cues, such as relationships between objects, familiarity with objects, concrete standards of spatial and body reference (in this case, for instance, peripheral vision of the nose for midline), and others. In a situation such as our own experiment, these cues are withheld, and consequently, there is an exaggerated operation of this compensatory mechanism which results, here, in displacements to the side opposite the body tilt.

If one considers these findings, within the framework of the developmental theories of Werner (9) or Piaget (4), one might say that orientation in space, in young children, is somewhat less determinate (or more labile) than corresponding perceptual performance in the adult. Werner would interpret these differences as a sign of increasing articulation of the "self" (or "body scheme") from the environment. Related to this increasing differentiation of body and object is the increasing capacity for the body to serve as a stable framework in perception. For the young child, the "exact middle of the head" apparently describes a broader and fuzzier area than for the adolescent or adult. Thus, the younger children tend to accept the auditory stimulus as given, and show far greater starting position effects than older Ss in Experiment I.¹

In a similar way, in Experiment II, the young child accepts a considerably tilted position as erect. For equal degrees of tilt, he appears to feel less tilted than the adolescent, and he does not have to be returned as far to the objective upright in order to feel upright. Just as "exact middle of the head" does not define a very precise body area for the child, so "sitting upright" does not define a very specific posture, at least when he has no visual cues.

The developmental evidence, so similar for audition and vision, points to some central rather than peripheral explanation for the increasing compensatory displacement of stimuli opposite the position of the body. Whatever these developing counteractive forces may be, it appears that they are the same whether the stimulus is to be localized in space or with reference to the self.

REFERENCES

1. AUBERT, H. Eine scheinbare bedeutende Drehung von Objekten bei Neigung des Kopfes nach rechts oder links. *Virchow's Arch. pathol. Anat. u. Physiol.*, 1861, 20, 381-393.
2. CARINI, L. An experimental investigation of perceptual behavior in schizophrenics. Unpublished doctoral dissertation, Clark Univer., 1955.
3. LIEBERT, R. S., WAPNER, S., & WERNER, H. Studies in the effects of lysergic acid diethylamide (LSD-25): visual perception of verti-

¹ Within each age group there are considerable individual differences which are reflected in the magnitude of perceptual "errors." Such factors have been explored by Witkin *et al.* (11) and are beyond the scope of this paper.

CHILD DEVELOPMENT

- cality in schizophrenic and normal adults. *Arch. Neurol. Psychiat.*, 1957, 77, 193-201.
4. PIAGET, J. *The origins of intelligence in children*. New York: International Universities Press, 1952.
5. TEUBER, H.-L., & LIEBERT, R. S. Effects of body tilts on auditory localization. *Amer. Psychologist*, 1956, 11, 430. (Abstract)
6. TEUBER, H.-L., & LIEBERT, R. S. Specific and general effects of brain injury in man: further evidence of their coexistence. *Arch. Neurol. Psychiat.*, 1958, 80, 403-407.
7. WAPNER, S., & WERNER, H. Experiments on sensory-tonic field theory of perception: V. Effect of body status on the kinesthetic perception of verticality. *J. exper. Psychol.*, 1952, 44, 126-131.
8. WAPNER, S., & WERNER, H. *Perceptual development: an investigation within the framework of sensory tonic field theory*. Worcester, Mass.: Clark Univer. Press, 1957.
9. WERNER, H. *Comparative psychology of mental development*. Chicago: Follett, 1948.
10. WERNER, H., & WAPNER, S. Experiments on sensory-tonic field theory of perception: IV. Effect of initial position of a rod on apparent verticality. *J. exper. Psychol.*, 1952, 44, 68-74.
11. WITKIN, H. A., & Associates. *Personality through perception*. New York: Harper, 1954.

PERCEIVED SIMILARITY TO PARENTS AND ADJUSTMENT

SUSAN W. GRAY*

George Peabody College for Teachers

This study represents an attempt to investigate certain relationships between perceived similarity to parents and indices of social and personal adjustment in a group of children in the fifth through eighth grades of a laboratory school.

Perceived similarity to parents is presumably one indication of parental identification. Although only one aspect of these processes which Freud subsumed under the term identification, it is clearly one such process. For example, Freud (7), in speaking of identification in very general terms states: "It may arise with every new perception of a common quality shared with some other person who is not an object of the sexual instinct" (p. 65). Perceived similarity also appears to be a way of tapping what Mowrer (16) has referred to as defensive identification, a process he contrasts with developmental identification. Stoke (25), in his careful analysis of the concept, makes a distinction between behavioral and emotional identification, the latter appearing to be a process revealed in part through the similarity one perceives between himself and a parent. There is some justification, then, for considering the present study as one of parental identification. As Knight (10) pointed out nearly two decades ago, however, no term in psychoanalytic literature has had more different usages than this one of identification. Certainly there would be ample cause to quarrel with the indices of perceived similarity presently to be described if they were taken to represent definitive and inclusive measures of parental identification.

Even if perceived similarity to parents is accepted as one way of tapping the process of identification, there is still the problem of appropriate indices to derive from data on parent-child similarities. A usual technique (4, 24)

* Box 232, George Peabody College, Nashville 12, Tennessee.

CHILD DEVELOPMENT

has been to use an inventory or check list which the subject marks as a self-report and then marks for his parent as perceived. The index of similarity then is based upon the number of items checked or rated in the same fashion for the self-report and for the parent as perceived. Certain problems arise with such a technique. One of these is that perceiving others as like oneself, or vice versa, is a generalized tendency, as Cronbach (5) has pointed out in another context. In the present study, for example, rank-difference correlations were .74 and .83 for boys and girls, respectively, between the measure of similarity to one parent and to the other. Because of this situation, the customary technique was not employed. Instead, two other indices of perceived similarity were used. The first of these was based upon the generalized perceived similarity to both parents, the argument being that the perceived closeness to or distance from both adults in the family group is a relevant variable in mediating adjustment. The second index is based upon the discrepancy between perceived similarity to parent of the same sex and parent of the opposite sex. It seems reasonable on a psychoanalytic basis to predict that an important aspect of perceived similarity to parents in relation to adjustment is whether the individual sees himself as more like his same-sex or his opposite-sex parent. The boy may see himself as very similar to both parents, but still as more like his father than his mother. Or he may see himself as very distant from both, but the same direction of perceived resemblance hold. The direction of perceived resemblance toward or away from a given parent would not on a *a priori* grounds appear related to an index of generalized similarity to both parents. The two indices were used to test different hypotheses in the present study.

Another problem inherent in the use of measures of perceived similarity, particularly in a study probably involving perceptions of sex roles as well as perceptions of specific individuals, is that of ambiguity. What does it mean that a given child sees himself and his father in a similar light? That this may be identifying with a sex-role rather than with a given parent is likely. Since Freud originally invoked the concept of identification in part, at least, to explain the adopting of an appropriate sex role, we would be in the familiar position of throwing out the baby with the bath water in attempting to untangle these two factors. A more serious difficulty is that of the relative esteem in which the child holds each of the sexes or the behavior commonly associated with each. If the child holds a more favorable picture of adults of his own sex, the degree of similarity which he perceives may reflect his own level of self-esteem. Masculinity is generally considered to be regarded more favorably in our culture; on the other hand, there is a piling up of evidence that young children show more regard for their mothers than their fathers. Some attempts were made to reckon with the ambiguity involved in the use of indices based upon perceived similarity to parents by the choice of the particular measures used in the present study and by the examination of the content of these instruments. These attempts to lessen the ambiguity of the measures will be discussed in subsequent sections devoted to the several instruments.

SUSAN W. GRAY

The hypotheses of this study stem originally from the psychoanalytic concept of parental identification. As Stoke (25), however, has pointed out, it is clear that certain variables may mediate patterns of identification. Selected for study in the present study were the variables of sex, age, and adjustment level. Three hypotheses were set up which involve these variables. These are given below, with a brief rationale for each.

1. In children of all grade groups combined there will be a positive relationship between sex-appropriate behavior as perceived by peers and identification directed toward parent of the same sex.

Freud used the concept of identification to explain how the individual learned to assume patterns of behavior perceived as appropriate to his sex. One need not, however, seek a psychoanalytic justification for this hypothesis. As Miller and Dollard (14) pointed out some years ago, children tend to be rewarded for imitation of the behavior of their same-sex parent. Thus, on the basis of secondary reinforcement one might expect the assuming of sex-appropriate behavior and the perceiving of oneself as more similar to the same-sex parent to go hand in hand.

2. There will be an interaction effect between sex and direction of identification for the entire group of children where indices of adjustment are considered. Boys who see themselves as more like their fathers will tend to show better adjustment; girls who see themselves as more like their mothers, however, will not show better adjustment.

This hypothesis is based in part upon the formulations of Erikson (6). In his writing on "Mom," Erikson has suggested that, while both boys and girls in America today may be expected to have problems in maintaining identification patterns with their same-sex parents, the girl's situation will be particularly difficult. The girl who identifies with her mother may be expected to have difficulties in personal relations. This line of reasoning is borne out in certain research which has compared the relationship of adjustment to parental identification in the male and the female. Thus, Helper (9), Lazowick (11), and Sopchak (24) have all found that, while identification with the father is related to adjustment in the male, identification with the mother is either unrelated or negatively related to adjustment in the female.

3. There will be an interaction effect between perceived distance from both parents and school grade, where the fifth and sixth grades combined are compared with the seventh and eighth grades. In the younger group adjustment will be associated with closeness to parents, while the reverse will be true in the older groups.

The older group may be expected to be passing into the phase of early adolescence in which achieving emancipation from parents becomes a paramount task. Textbook treatments of adolescence stress this problem as well as do general writings upon developmental tasks in our culture. Rejection of parental values and hence perceived dissimilarity to parents may be expected to be more characteristic of those seventh and eighth grade youngsters making a favorable adjustment than it would be of the well-adjusted

CHILD DEVELOPMENT

youngsters of the earlier grades. The younger children are still rather closely tied to the family group and more emotionally dependent upon it.

METHOD

Instruments

Index of identification. The instrument used in the present study to assess parental identification is a list of 40 bipolar adjectives or phrases modeled after Helper (9) but simplified for use with young children. This list is roughly descriptive of the factors which Cattell (3) found in his factor analysis of personality ratings. Each adjective pair, following Osgood's format (17), is rated on a 7-point scale. Illustrations of the list are:

brave								timid
tired								full of pep
easy-going								serious-minded

This list was used both for the child's self-report and for the parent as perceived by the child.

Two indices of perceived similarity were derived from this list. The *D* statistic of Osgood and Suci (18) was computed for each child and his perceived father, and similarly for the child and his perceived mother. The index of distance from parents was simply the sum of the *D*s for the child—perceived father and child—perceived mother comparisons. For the index of direction of identification use was made of the discrepancy between the *D*s for the two parents. These measures are admittedly crude. They were used in the study solely for the purpose of constituting extreme high and low groups upon the index in question.

Measures of adjustment. Two measures of adjustment were employed. The first of these was the *Who Are They* test, a reputation test developed for use in the Community Youth Development Project (1) of the University of Chicago. With this sociometric test children are asked to nominate their classmates for 15 items designed to tap five behavioral variables: practical intelligence, leadership, friendship, withdrawal, and aggression. Actually these five variables are not independent, although based upon separate items on the test. Mitchell (15) in a factor analysis of the test found evidence for three factors. The first of these was a generalized social acceptance, made up chiefly of the first three variables. The other two were a socially isolated and an aggressive maladjustment, corresponding in general to the withdrawal and aggressive variables. The five variables were retained for the present study, however, for the purpose of giving some picture of the consistency of the relationships between perceived similarity and social acceptance or rejection. In common with other reputation tests, the reliability of the test appears high, except for the aggression variable in girls (8). Intercorrelations on the five variables for the present sample tended to be

SUSAN W. GRAY

similar to those found by Mitchell. In making their nominations for each item, children were given unlimited choice. A question arises as to the sex bias in this instrument. In an effort to check for this, comparisons were made of the mean number of nominations given each sex on each of the five variables. These means are given below.

	Practical Intelligence	Leadership	Withdrawal	Aggression	Friendship
Boys	10.5	11.9	7.0	7.3	4.8
Girls	10.5	11.7	7.6	5.8	4.5

It appears that only on the aggression variable is there any appreciable tendency to nominate one sex more frequently than the other.

The second index of adjustment used was the children's form of the manifest anxiety scale (2). This test has been shown to have some relation to a sociometric variable (13), and to school achievement (12). It seems reasonable that scores upon this test might be associated with difficulties in meeting the developmental expectancies of one's age and sex group, particularly as related to the parent-child situation. Thus, it would seem relevant here as an index of adjustment.

The masculinity-femininity scale. This scale, devised by the writer for a previous study (8), is a 5-point rating scale, with descriptions for each point, on which each child is asked to rank every other child in the group. The sexes are rated separately. The score obtained is the mean of all the rankings an individual receives from his peers. In the earlier study mentioned above, test-retest coefficients of correlation for sixth and seventh grade youngsters were .86 ($N = 34$) for boys and .94 ($N = 27$) for girls over a six-week period. In the present study, reliability checks were run only with the fifth grade. Here test-retest coefficients of correlation were .87 ($N = 14$) for boys and .95 ($N = 16$) for girls over a period of three weeks.

Subjects

All children in the fifth to eighth grades inclusive of a laboratory school were tested with the instruments described above. These children were distributed as follows by grade and sex:

	Fifth	Sixth	Seventh	Eighth
Boys	11	18	21	16
Girls	15	7	11	16

For purposes of analysis, the fifth and sixth grade groups were combined, and in a like fashion the seventh and eighth grade groups.

Two limitations concerning the sample should be mentioned. The major one is the use of a laboratory school. Typically these children tend to come from middle class homes, particularly the upper middle class.

CHILD DEVELOPMENT

Social class has been shown to be related to peer status (19). It also seems reasonable that differences in parental identification should appear, not only on the basis of child rearing practices, but also because of the role model provided by the upper middle class father and mother. A second limitation is the uneven sex distribution of the sixth and seventh grades. The grouping together of the fifth and sixth, and of seventh and eighth, was used in an effort to offset the possible influence of this situation.

In the analyses approximately the top and bottom one-thirds of each sex for the younger group and a similar number for the older group for the perceived similarity variables were used. This was done separately for the distance and direction indices.

The children were tested in class by their teachers, each of whom had been given an opportunity to become familiar with the instruments. An attempt was made to space the tests so as to break up any general sets which might be formed.

RESULTS

Descriptive Data on the Instruments Used

The adjective check list. With a 7-point scale and with 40 items, the *D* statistic for this instrument has a theoretical range of 0 to 37.8 for each parent singly or 0 to 75.6 for the sum of the *D*s for the two parents. Ranges and medians for the *D*s are given in Table 1. The medians tend to be similar for the age and sex groups. Of interest on the directional variable is the shift in boys from the lower to the upper grade group in the direction of similarity. Of the 26 in the younger group, only five saw themselves as more like their fathers, while of the 34 in the older group, 18 saw themselves in this wise. The girls showed little shift, nine of the 19 in the younger group seeing themselves as more like their mothers, and 16 of the 27 in the older group. The shift in the boys is reasonable enough from the standpoint of a pulling away from a mother-dominated pattern of behavior.

TABLE 1
MEDIAN AND RANGES FOR THE *D* MEASURE OF PERCEIVED
SIMILARITY TO PARENTS

	PERCEIVED SIMILARITY TO FATHER		PERCEIVED SIMILARITY TO MOTHER		SUM OF THE TWO MEASURES	
	Median	Range	Median	Range	Median	Range
Boys of Grades 5 and 6 ..	12.6	5.3-18.5	11.4	5.6-14.8	23.0	10.9-33.3
Boys of Grades 7 and 8 ..	11.1	5.7-24.6	11.4	5.8-23.7	22.4	12.7-48.3
Girls of Grades 5 and 6 ..	11.5	5.6-22.5	10.7	6.4-21.6	22.0	12.1-44.1
Girls of Grades 7 and 8 ..	10.6	5.0-17.3	9.5	5.4-16.2	20.6	11.5-31.5

SUSAN W. GRAY

It was surprising, however, that so few of the younger boys saw themselves as more like their fathers. It remains to be seen whether this is a function of the particular sample or has a wider generality.

Inspection of the ranks given to fathers and mothers do not indicate any appreciable sex bias in the extent to which one sex is more favorably rated than the other by the children of the present sample. Of the 40 bipolar adjective pairs, 31 were considered by a jury of 12 graduate students in psychology to be to some extent on a continuum from favorable to unfavorable. Illustrations are: brave—timid, happy—sad, and lazy—hard-working. Nine pairs were considered as not being clearly on such a continuum, as easy-going—serious-minded and heavy—light. On the 31 favorable-unfavorable items, the mothers were seen more favorably on 19 items, and the fathers on 12, for the 49 boys in the sample. For the 66 girls, the mothers were seen more favorably on 18 of the items and the fathers on 13. A sign test with the boys' data indicated that such differences as appeared in the number of items in which one sex is seen more favorably than the other might be expected to occur by chance approximately 28 per cent of the time. With girls, of course, the probability of such a difference's occurring by chance is even greater.

The masculinity-femininity rating scale. The mean ratings and their standard deviations for the index of sex-appropriate behavior are given below:

	<i>M</i>	<i>SD</i>
Fifth and sixth grade boys	2.47	.47
Fifth and sixth grade girls	2.75	.57
Seventh and eighth grade boys	2.48	.55
Seventh and eighth grade girls	2.59	.54

Differences in the means of the groups were not significant. Since in every case the lower score refers to the more sex-appropriate behavior, it may be seen that on this scale from 1 to 5 the group tended to rate their peers somewhat more toward the sex-appropriate end of the continuum.

The children's form of the manifest anxiety scale. Means and standard deviations for this group by grade and sex were as follows:

	<i>M</i>	<i>SD</i>
Fifth and sixth grade boys	13.4	5.3
Fifth and sixth grade girls	16.9	9.8
Seventh and eighth grade boys	14.8	6.1
Seventh and eighth grade girls	15.6	7.9

The authors of the test (2) quote means of 16.6 and 18.5 for boys and girls, respectively, of the sixth grade. Corresponding standard deviations were 7.4 and 7.8. Fifth grade scores were similar. For the higher grades norms

CHILD DEVELOPMENT

are not given. The present sample, then, has slightly lower scores than the normative group in the grades for which we have data. In the analysis of data all children were eliminated who, on the lie-scale incorporated within the test, scored higher than five.

The Who Are They test. As might be expected, scores on the reputation test varied widely. There were a large number of 0 scores on every category, particularly with the withdrawal and aggression dimensions. The scores ranged upward in the J-curve typical of sociometric data.

Statistical Tests of the Hypotheses

Since the distributions of scores on the reputation test were markedly skewed, a nonparametric test, the Mann-Whitney U test (23), was employed. For testing the interaction hypotheses the data were arranged in a 2×2 table and a slight modification of the Mann-Whitney test was employed. An illustration is given in the next paragraph of how this was done with one variable used in testing an hypothesis involving an interaction.

According to the second hypothesis, we would predict that, on the anxiety variable, boys high in identification with the same-sex parent would have low scores on the anxiety test, while girls high in identification with the same-sex parent would have high scores on the anxiety test. To test this interaction, and at the same time make possible tests of the main effects for sex and for level of identification, the data are grouped as shown in the diagram below with H indicating high identification with same-sex parent, L, low identification, 5, the fifth-sixth grade grouping, and 7, the seventh-eighth grade grouping. These symbols will be used in the tables of the paper. The cell entries in the analysis are scores upon the anxiety test.

L5B	L5G
L7B	L7G
H5B	H5G
H7B	H7G

If the cell entries are proportional, the predicted interaction may be tested in this 2×2 table by summing the diagonals and using these two sums for a conventional U test. Thus, if the sum for the diagonal from upper left to lower right is enough greater than the opposite diagonal that U is statistically significant, the prediction is confirmed. A U test may be made of the main effect for sex by summing the columns, and a test of the main

effect for level of identification by summing the rows. A test of interaction for grade by level of identification may be made by rearranging the subgroups so that a cell consists of a grade group rather than a sex group. It should be pointed out that the U tests of rows, columns, and interaction possible in the sort of analysis described are not independent. A large row or column effect may possibly serve to decrease the power of the Mann-Whitney U as a test of interaction.

When the analog with analysis of variance is pushed further, and one attempts to make tests of simple effects, another problem, involving loss of power enters in. Here, one has recourse only to the U test with the smaller number of cases involved in testing the simple effects. In other words, one does not have the advantage of the reduced error term possible in *t* tests of simple effects where analysis of variance is used. This limitation and the one discussed in the preceding paragraph probably increase the likelihood of a Type II error in the analyses which follow.

The data of the present study, particularly the *Who Are They* scores, involved a large number of ties. For this reason the correction for ties given by Siegel (23) was used for each of the Mann-Whitney U tests. The .05 level of significance was set, and since all hypotheses were directional in nature, one-tail tests were used.

The first hypothesis. Here it was predicted that children perceiving themselves as more similar to their same-sex parent would be seen by their peers as more sex-appropriate in behavior than those children seeing themselves as less like their same-sex parent. This prediction was not confirmed. A $U - \mu$ of 123 in the predicted direction was found. This, however, is only 1.34 times σ_U . Instead, a significant nonpredicted grade by direction of identification interaction was obtained. Here $U - \mu = 313$, its ratio to σ_U being 3.39. All the tests for simple effects for this comparison were significant at or beyond the .05 level for a one-tail test. This finding seems contrary to reasonable prediction. It may be that perceptions of sex role do undergo some major reorganization during this age group. More likely perhaps is the chance that this results from some peculiarities of the present sample. These children came predominantly from upper middle class homes. Many of their fathers were professional men. College teachers and physicians may not be as masculine in the eyes of the young as truck drivers and construction workers. It is possible that the particular sample of laboratory school mothers of the younger group were more striving and dominant than women of the general population. Thus, a child might be identifying with his mother and at the same time with the more masculine figure in his household, as masculinity is defined by his peers.

Another factor which may have led to the nonconfirmation of the prediction is that in the younger children the majority, male or female, saw themselves as more like their mothers; they simply saw themselves as less like their mothers than did the low group. Because of this situation, it was decided to do some *post hoc* hypothesizing and test the first hypothesis for the seventh and eighth grades separately. When high and low

CHILD DEVELOPMENT

identification groups were compared for the older children on the masculinity-femininity dimension, $U - \mu$ in the predicted direction was equal to 112. Its ratio to σ_U was 2.81, significant at the .05 level. It may be, then, that the failure to confirm the hypothesis for the total group results from the unsuitability of the directional index for the younger children.

The second hypothesis. The general prediction here was that boys who perceived themselves as more similar to their fathers than their mothers would show better adjustment; girls who perceived themselves as more similar to their mothers would not show better adjustment. An interaction effect is thus expected between sex and the directional variable for the five

TABLE 2
MEDIAN SCORES ON THE ADJUSTMENT INDICES FOR THE
DIRECTIONAL VARIABLE

Adjustment Variable	Low Boys			Low Girls			High Boys			High Girls		
	5th	7th	T	5th	7th	T	5th	7th	T	5th	7th	T
Practical Intelligence ..	2	2	2	6	15	12	17	10	12	6	7	7
Leadership ...*	2	0	1	8	14	10	12	6	9	12	5	8
Friendship	4	2	3	1	8	6	5	5	5	5	2	4
Withdrawal	6	10	7	6	5	5	1	4	1	2	9	2
Aggression	4	5	4	1	6	4	0	2	2	4	2	4
Manifest Anxiety*	15	15	15	15	14	15	12	15	14	15	15	15
N	9	12	21	7	9	16	9	12	21	7	9	16

* For manifest anxiety, Ns should read in order across the table: 8, 11, 19; 7, 8, 15; 8, 11, 19; 7, 8, 15.

categories of the *Who Are They* test and for the anxiety scale. Table 2 shows the medians obtained on the *Who Are They* test and the anxiety scale for boys and girls, as divided into groups designated as high or low in terms of discrepancy between the *D* statistic for the same-sex parent and for the opposite-sex parent. High indicates a relatively greater resemblance to the same-sex parent, and low, a relatively less resemblance. In Table 3 are shown the results of the analysis performed to test the prediction of interaction between sex and the directional variable. Predictions were confirmed in four of the six comparisons. The other two, while in the predicted direction, fell short of significance. The reader should probably be reminded here that the five categories of the *Who Are They* test are independent only in the sense that they are derived from different items on the test. Mitchell's (15) analysis would indicate that the variables of practical intelligence and friendship, for both of which positive results were obtained, are tapping a single factor, that of social effectiveness.

When the simple effects are examined, it is in the comparisons involving the boys low in identification with same-sex parent that significant results

SUSAN W. GRAY

are obtained. All four comparisons of high and low boys' results were significant at the .05 level; the only other two significant comparisons were ones involving low boys and low girls. All other comparisons were in the predicted direction, although not significant.

Tests of interaction for grade by direction of identification were also made, although such relationships had not been predicted. None of the over-all tests proved significant.

At first blush it seems difficult to see why the prediction of an interaction between sex and direction of identification should have been confirmed in the majority of cases, while that concerning masculinity-femininity and

TABLE 3
TESTS OF SIGNIFICANCE FOR THE DIRECTIONAL VARIABLE

<i>Adjustment Variable</i>	<i>Over-all Test of Predicted Direction by Sex Interaction</i>	<i>Simple Effects</i>			
	<i>N = 74</i>	<i>HB—HG N = 37</i>	<i>LB—LG N = 37</i>	<i>LB—HB N = 42</i>	<i>LG—HG N = 32</i>
Practical Intelligence ..	2.32*	.95	1.87*	2.91*	.25
Leadership	2.83*	1.20	2.43*	3.06*	.95
Friendship	2.22*	1.60	1.19	1.73*	1.11
Withdrawal	1.89*	1.57	.98	2.57*	.93
Aggression	1.43				
Manifest Anxiety†91				

* Significant at the .05 level for a one-tail test.

† For manifest anxiety, *N* = 70.

direction of identification was not confirmed. Sex-appropriate behavior as measured in this study has been found to be positively related to adjustment in boys, and to a lesser extent in girls (8). Perhaps we should here make a distinction between identification with one's own father and identification with a masculine figure. Identification with the particular fathers of the younger age group of this sample possibly has not led to sex-appropriate behavior but has led to adjustment as perceived by one's peers.

The question arises as to the failure to confirm the hypothesis with the aggression variable. This variable appears to be heterogeneous, which may explain the low degree of relationship. Also, as mentioned earlier, the aggression variable has the lowest reliability of the five reputation test variables. It was also the only one of the five variables on which was found an appreciable difference in the mean number of nominations for the two sexes.

The failure to confirm the hypothesis concerning the anxiety variable is less easy to explain. Unless this sample is atypical in this respect—an obvious possibility—the hypothesis may be incorrect, or else too many uncontrolled sources of manifest anxiety are at work.

CHILD DEVELOPMENT

The third hypothesis. Here it was hypothesized that in the lower grades children who saw themselves as closer to their parents would be seen by their peers as better adjusted than those who saw themselves as distant from their parents; in the upper grades the trend would be reversed. Thus, an interaction between grade and distance from parent is expected on the *Who Are They* indices and on the anxiety scale. Table 4 gives the medians obtained on the *Who Are They* indices and on the anxiety scale for boys and girls, divided into low and high groups on the basis of their closeness or distance, respectively, from their parents. In Table 5 are shown the results of the analysis of predictions based on the hypothesis of an interaction between grade level and distance from parent. With regard to the over-all tests of interaction, the predictions are confirmed in four of the six instances. The simple effects, however, as with the directional variable, are significant in less than half the cases. Three of the six comparisons on the *Who Are They* variables involving the upper grade group more distant from their parents are significant; the only other two significant comparisons are with the anxiety variable. In this table, as in Table 3, all other comparisons are in the predicted direction, although not significant. In general, it appears that seeing oneself as distant from one's parent is a correlate of social acceptance with the older children. Its influence with the younger children is less clear-cut.

TABLE 4
MEDIAN SCORES ON THE ADJUSTMENT INDICES FOR THE
DISTANCE VARIABLE

Adjustment Variable	Low 5th Grade			Low 7th Grade			High 5th Grade			High 7th Grade		
	B	G	T	B	G	T	B	G	T	B	G	T
Practical Intelligence . .	6	18	10	10	20	10	6	2	4	10	15	15
Leadership	9	12	10	2	5	3	4	5	5	10	12	12
Friendship	5	5	5	4	4	4	4	3	4	6	6	5
Withdrawal	2	2	2	5	5	5	1	11	4	4	6	6
Aggression	2	1	1	4	2	4	5	7	6	7	3	4
Manifest Anxiety* . . .	12	12	12	18	13	16	15	25	16	13	23	17
N	9	7	16	12	9	21	9	7	16	12	9	21

* For manifest anxiety, Ns should read in order across the table: 8, 7, 15; 11, 8, 19; 8, 7, 15; 11, 8, 19.

Tests of interaction for sex by the distance variable were also performed, although no such relationships had been predicted. The largest z was that for withdrawal. Here $z = 1.90$. This is not significant at the .05 level for a two-tail test. Since this was a nonpredicted interaction, the ratio was considered too low to be statistically significant. Such relationship as exists,

SUSAN W. GRAY

however, is in the direction of withdrawal's being characteristic of boys who see themselves as like their parents, and of girls who see themselves as distant from their parents.

The relationship discussed in the paragraph above may be a partial explanation of the failure to confirm the predicted grade by distance interaction where withdrawal is concerned. This explanation is given credence by the study of Schoeppe (22), who found that autonomy—and thus presumably seeing oneself as dissimilar to parents—was more important for adolescent boys than for girls.

The speculations concerning the failure of the aggression variable to show a significant relationship with the direction of perceived similarity would appear to apply here as well since the comments were concerned with the general inconsistency and presumed heterogeneity of the aggression index.

From the standpoint of the over-all tests, the hypothesis is confirmed for those variables dealing with social acceptance and with anxiety. The results of the tests of simple effects, so-called, however, indicate that the interaction is either not strong enough or else not consistent enough to appear in more than five of the 16 tests. It probably indicates also that the use of the Mann-Whitney U test for comparisons of simple effects is less efficient than analogous comparisons with parametric statistics.

TABLE 5
TESTS OF SIGNIFICANCE FOR THE DISTANCE VARIABLE

Adjustment Variable	Over-all Tests of Predicted Distance by Grade Interaction N = 74	Simple Effects			
		H7—L7 N = 42	H5—L5 N = 32	H5—H7 N = 37	L5—L7 N = 37
Practical Intelligence ..	1.70*	1.21	1.49	1.84*	.37
Leadership	1.90*	1.53	1.32	1.09	1.39
Friendship	1.93*	2.23*	.64	1.79*	.76
Withdrawal49				
Aggression76				
Manifest Anxiety†	1.97*	.75	2.06*	.96	2.02*

* Significant at the .05 level for a one-tail test.

† For manifest anxiety, Ns should read in order across the tables: 68, 38, 30, 34, 34.

DISCUSSION

The general findings of this study would seem to indicate that perceived similarity to parents does have correlates in behavior relating to adjustment. Boys who perceive themselves as more like their fathers than their mothers

CHILD DEVELOPMENT

are perceived more favorably by their peers. With girls, to the extent that a relationship exists, it is in the direction of girls who see themselves as more like their mothers being seen less favorably by their peers. The predictions of this study have to a limited extent, then, been borne out. The rationale for these hypotheses, it will be remembered, stems from changes in role expectancies for women today and the general devaluation of feminine characteristics which Tuddenham (27) found to characterize the youngsters in the upper elementary grades. An extrapolation of the findings of Tuddenham and the earlier ones of Tryon (26) is usually made to the effect that as girls pass beyond the age of 14 or 15 the more feminine traits will again have high valuation. Hence, we would expect that identification with the mother or at least a feminine role model would be related to social acceptance for the girl during later adolescence. An empirical test of this predicted change would seem a valuable next step in a study of parental identification as related to adjustment.

The findings concerning distance from parents support the hypothesis concerning this variable in two-thirds of the over-all tests. It appears to be important to reject perceived similarity to parents as one goes from the younger to the older groups of this study. This is shown both in the social acceptance indices and in the anxiety test. Here again is a finding which one could profitably compare with results from a yet older group. There is some evidence that perceived dissimilarity to others is, in the adult, indicative not of adjustment but of disorder (21). It might be reasonable to predict that when these youngsters develop beyond the stage of "protesting too much" they may again see themselves as close to their parents.

At least two issues are left unresolved by the present attack upon parental identification. One of these relates to the nature of the sample used. Results can probably be generalized to a population of laboratory schools and with caution to an upper middle class group. Since there is some evidence that sex-role identification differs in the social classes (19) and also that peer acceptance is based upon somewhat different characteristics in lower and middle class groups (20), it would seem desirable to test the hypotheses of the current study with other social class groupings. Particularly relevant would be a test of the hypothesis concerning direction of identification. In groups where fathers and mothers might be expected to follow somewhat more clearly differentiated patterns of masculine and feminine behavior, the predictions concerning direction of identification might be confirmed to an even greater extent.

The other issue is that of developing adequate measures of identification. The indices of perceived similarity used in the present study are too limited and to some extent too ambiguous to give more than a partial answer to questions of the relation of parental identification to adjustment. More refined and less equivocal measures are needed before the variable of parental identification, so often employed by personality theorists as an explanatory principle, can be adequately assessed.

SUSAN W. GRAY

SUMMARY

Relationships between perceived similarity to parents and indices of personal and social adjustment were studied in a group of children in the fifth through eighth grades of a laboratory school. Perceived similarity was considered to be indicative of parental identification.

Two indices of identification were used: the direction of identification toward or away from the same-sex parent, and the degree of distance from both parents combined. Analyses were performed by means of a modified Mann-Whitney U test. It was predicted that there would be an interaction effect between direction of identification and sex, identification with the father being positively associated with adjustment in the boy, while with the girl identification with the mother would not be associated with adjustment. A significant interaction was obtained in four of the six over-all tests. Tests of simple effects indicated that the direction of identification was a more important factor for the boy than the girl with respect to adjustment.

It was further predicted that there would be an interaction between age level and distance or closeness to parent. Four of the six over-all tests of interaction proved significant. Examination of simple effects, however, revealed significant comparisons in only five of the 16 possible tests. Pattern is not consistent as to where the significant comparisons occur.

The study is interpreted as indicating the importance of the changed role expectancies for women today and of the pressure toward emancipation from parents in younger adolescents as mediating variables in determining patterns of identification in the age groups studied. Further research is needed to see if the relationships here demonstrated with an upper middle class group hold in other social classes and in a wider age range.

REFERENCES

1. BOWMAN, P. H., *et al.* Studying children and training counselors in a community program. *Suppl. Educ. Monogr.*, 1953, No. 38.
2. CASTANEDA, A., McCANDLESS, B. R., & PALERMO, D. The children's form of the manifest anxiety scale. *Child Developm.*, 1956, 27, 317-326.
3. CATTELL, R. B. *Personality: a systematic, theoretical and factual study.* New York: McGraw-Hill, 1950.
4. CAVA, E. L., & RAUSCH, H. L. Identification and the adolescent boy's perception of his father. *J. abnorm. soc. Psychol.*, 1952, 47, 855-856.
5. CRONBACH, L. J. Processes affecting scores on "understanding of others" and "assumed similarity." *Psychol. Bull.*, 1955, 52, 177-193.

CHILD DEVELOPMENT

6. ERIKSON, E. H. *Childhood and society*. New York: Norton, 1950.
7. FREUD, S. *Group psychology and the analysis of the ego*. New York: Boni & Liveright, 1922.
8. GRAY, SUSAN W. Masculinity-femininity in relation to anxiety and social acceptance. *Child Develpm.*, 1957, 28, 203-214.
9. HELPER, M. M. Learning theory and self-concept. *J. abnorm. soc. Psychol.*, 1955, 51, 184-194.
10. KNIGHT, R. P. Introjection, projection, and identification. *Psychoanal. Quart.*, 1940, 9, 334-341.
11. LAZOWICK, L. M. On the nature of identification. *J. abnorm. soc. Psychol.*, 1955, 51, 175-183.
12. McCANDLESS, B. R., & CASTANEDA, A. Anxiety in children, school achievement and intelligence. *Child Develpm.*, 1956, 27, 379-382.
13. McCANDLESS, B. R., CASTANEDA, A., & PALERMO, D. Anxiety in children and social status. *Child Develpm.*, 1956, 27, 385-392.
14. MILLER, N. E., & DOLLARD, J. *Social learning and imitation*. New Haven: Yale Univer. Press, 1941.
15. MITCHELL, J. V., JR. The factor analysis of a "Guess Who" questionnaire designed to identify significant behavior patterns in children. *J. Pers.*, 1956, 24, 376-386.
16. MOWRER, O. H. *Learning theory and personality dynamics*. New York: Ronald, 1950.
17. OSGOOD, C. E. The nature and measurement of meaning. *Psychol. Bull.*, 1952, 49, 251-262.
18. OSGOOD, C. E., & SUCI, G. J. A measure of relation determined by both mean difference and profile information. *Psychol. Bull.*, 1952, 49, 251-262.
19. POPE, B. Socio-economic contrasts in children's peer culture prestige values. *Genet. Psychol. Monogr.*, 1953, 48, 157-220.
20. RABBAN, M. Sex-role identification in young children in two diverse social groups. *Genet. Psychol. Monogr.*, 1950, 42, 81-158.
21. RUDIKOFF, ESSELYN C. A comparative study of the changes in the concepts of the self, the ordinary person, and the ideal in eight cases. In C. R. Rogers & Rosalind F. Dymond (Eds.), *Psychotherapy and personality change*. Chicago: Univer. Chicago Press, 1954. Pp. 85-98.
22. SCHOEPPPE, AILEEN. Sex differences in adolescent socialization. *J. soc. Psychol.*, 1953, 38, 175-185.
23. SIEGEL, S. *Nonparametric statistics for the behavioral sciences*. New York: McGraw-Hill, 1956.

SUSAN W. GRAY

24. SOPCHAK, A. L. "Parental identification" and tendency toward disorders as measured by the MMPI. *J. abnorm. soc. Psychol.*, 1952, 57, 159-165.
25. STOKE, S. M. An inquiry into the concept of identification. *J. genet. Psychol.*, 1950, 67, 162-189.
26. TRYON, CAROLINE M. Evaluations of adolescent personality by adolescents. *Monogr. Soc. Res. Child Developm.*, 1939, 4, No. 4.
27. TUDDENHAM, R. D. Studies in reputation. III. Correlates of popularity among elementary-school children. *J. educ. Psychol.*, 1951, 42, 257-276.

THE RELATIONSHIP BETWEEN CHILDREN'S EXPRESSIONS OF HOSTILITY AND METHODS OF DISCIPLINE EXERCISED BY DOMINANT OVERPROTECTIVE PARENTS

CHARLES Y. NAKAMURA*

University of California

In recent years, several experimenters have applied Miller's conflict and displacement theories (2, 10) to the study of children's aggressive behavior (4, 8, 13). A frequent finding has been the relation between aggressive activity and the methods of discipline used by the parents. Sears (12) found that children who directed the most aggression toward parent figures in doll play came from homes where punishment was severest. Hollenberg and Sperry (4) reported similar results. In a study by Bandura and Walters (1) it was shown that the occurrence of aggressive behavior among delinquents was associated with frequent physical punishment by parents.

The present study seeks to apply Miller's displacement and conflict paradigms in the further examination of some observations, made in a counseling situation, of methods of discipline and expression of hostility. In the process of counseling high school and college students for personal problems, it was noted that a large number of them, especially girls, had great difficulty in expressing negative feelings and criticism of their parents. This obtained in spite of the fact that parent-child interactions that would justify feelings of resentment on the part of the child were repeatedly described. That resentment was indeed experienced was verified by eventual verbalization of it in later phases of counseling. These counselees characteristically described their parents as dominant and overprotective. However, this was done indirectly in the context of descriptions of much love within the family where parents had considerable concern for the welfare and happiness of their children. There was a strong sense of obligation which served the means by which the parents imposed strong controls. Thus, even when the control sharply restricted the children's independence to an extent where they might reasonably experience resentment, the sense of obligation inhibited such negative feelings.

* Department of Psychology, University of California, Los Angeles 24.

CHILD DEVELOPMENT

According to Miller's conflict and displacement hypotheses, it would be expected that under such conditions of dominant overprotectiveness, the children would have to learn to express these negative feelings in directions other than toward their parents. This follows from the principle that when the direct response to the original stimulus object is prevented by a conflict, it will be directed toward another object sufficiently dissimilar from the original one to reduce the conflict.

This group of counselees was distinguishable from a second group who also described dominant overprotective parents but who were able to express criticism of them, although not without accompanying feelings of guilt. These students could be further differentiated from the former group in that their parents frequently administered more obviously severe punishments such as physical and verbal whippings. It seems that these counselees were able to express criticism of their parents because it was justifiable in light of the severe punishment which was administered. Consequently, conflict between expressing or not expressing resentment was less intense than it would be for the former group whose parents used subtle means of control which did not warrant criticism. This follows from the principle of conflict and displacement that decreasing the conflict shifts the displacement of aggression in the direction of the stimuli that are more similar to the original one. Thus the lower the intensity of the conflict, the more likely that the response will be directed toward the original object, the parents.

The conceptualization of the behavior of these counselees in terms of conflict and displacement led to the formulation of the hypotheses to be tested in this study.

1. Subjects whose parents exercised positive types of discipline report more favorable attitudes toward their parents than do subjects who are disciplined by negative methods.
2. Of subjects who have dominant overprotective parents, those disciplined by positive methods displace criticism to persons other than their parents more than do those disciplined by negative methods.
3. Of subjects whose parents did not use overprotective techniques of child rearing, there is no difference between those subjects disciplined by positive and those disciplined by negative methods in the amount of criticism displaced toward persons other than their parents.
4. Overprotected and positively disciplined subjects displace more criticism toward persons other than their parents than do nonoverprotected and positively disciplined subjects.

METHOD

Subjects

The subjects were 78 women in an introductory course in the psychology of adjustment. Due to the nature of the course, it may have included a greater proportion of students concerned with personal problems than most college courses.

Materials and Scoring

Measures were obtained on: (a) method of discipline exercised by parents, (b) parents' dominant overprotective or nonoverprotective techniques of child rearing, (c) children's criticism of parents and family, and (d) children's criticism of people other than their parents.

The measures were based on the responses to questions selected from an autobiographical workbook designed for use in college introductory courses in the psychology of adjustment (6). The book contains specific questions on topics such as family background, family relationships, social adjustments, and adjustments to frustration and to conflicts between goals. It is intended to facilitate efforts of the student to understand his own behavior and experiences.

Discipline. The question selected to measure method of discipline was: Discuss types of discipline used, how administered, by whom, how frequently, and your reactions to it at the time. Would you discipline your own child in the same way? Why or why not? Responses were scored positive or negative. There was no differentiation made as to which parent administered what type of discipline. The rating was based on a global judgment determined by type, consistency, frequency, and intensity of the disciplining methods. (The primary determinant for a rating of positive discipline was that the subject stated that she would discipline her child in the same way and gave reasons for this. A negative rating was determined by the frequency and intensity of physical punishment and the subject's statement of dissatisfaction with it.)

Criticism of parents. A score of positive or negative (critical) attitude toward the parents and intrafamily relations was determined by responses to questions on: (a) Family description—each parent was described on appearance, personality, special assets, special liabilities, characteristics liked most, characteristics disliked most. (b) Family happiness—was your family life a happy experience? Indicate specifically what made it so, or what made it an unhappy experience. Were your mother and father happy? Why or why not? (c) Special problems—did your family have special problems . . . ? The score of positive or negative criticism of parents was based on a global judgment including responses to all three of the questions.

Dominant overprotectiveness. Only dominant overprotection was considered in the study. Subjects reporting indulgent overprotection were not included. Two questions were selected for the measure of overprotected or nonoverprotected: (a) Do you believe your parents overprotected you? If so, describe their behavior upon which you base these feelings. Indicate the kind of overprotection most frequently employed—for example, indulgent overprotection or dominating overprotection, etc. (b) Indicate the efforts your parents made to teach you responsibility and to encourage your independence—or their efforts in the opposite direction. Have you wanted more independence? (The crucial determinant for a rating of dominant

CHILD DEVELOPMENT

overprotection was the presence of restriction of independence accompanied by implementation of control through parental overconcern for the welfare of the child, and that most children or adolescents subjected to such restriction would experience frustration and have justifiable feelings of resentment.) The following are excerpts from responses that were rated as overprotected:

I believe my parents overprotected me, at least more so than my friends as I was constantly restricted to home when I wanted to attend dances. Even in high school I was prohibited to go on trips until a senior. It was dominating overprotection, consequently I remained close to home, rarely left home during the summer during the day.

... as long as I remember they would never let me do or go anywhere I wanted.

I believe that my parents often tried to overprotect me. Mother especially tried to keep me from meeting hard knocks and disappointments by trying to fight my battles with friends for me. She was also especially overprotective when I got in high school by trying to keep me away from situations such as staying out late to parties or going steady with boys. . . . This caused a lot of trouble.

... as much as my parents wanted me to be able to shoulder responsibility they balked against my move for independence. Both mother and father tended to overprotection and always wanted to have a say in all of my decisions and were inclined to try to hold me back from each innovation as I grew up.

Criticism of others. This was selected as the criterion measure of displacement on the assumption that, if expression of resentment toward parents is inhibited by conflict, criticism of others would be a displacement of such feelings. High or low criticism of others was based on a section of the workbook entitled "Immature ways of reacting to conflict." It included these instructions: "The following are types of reaction commonly observed when persons are frustrated or in conflict. For each, provide a detailed description of something you have done, or seen others do which illustrates this type." The types of reactions are those commonly referred to as defense mechanisms such as rationalization, negativism, regression, compulsions, etc. There were 12 in all. A subject was rated high in criticism of others if eight or more of her examples of immature ways of reacting referred to people other than her parents or herself, and she made highly critical or condemnatory statements in regard to their immature behavior.

Procedure

Every member of the class completed the autobiographical workbook as a course requirement. The objectives of the course, protection of anonymity by assignment of a code number, and complete assurance that only the instructor and a specific graduate student reader would read the work maximized the frankness and honesty with which the questions were answered. The responses to the questions selected to measure each variable.

CHARLES Y. NAKAMURA

were independently rated by two raters to establish scoring reliability.¹ All scores were dichotomous.

The subjects were divided into two groups depending on whether they scored high or low on the measure of dominant overprotection by parents. They are referred to as the *overprotected* group ($N = 31$) and the *non-overprotected* group ($N = 47$). The relation between type of discipline exercised by parents and criticism of parents were analyzed for each of these groups. The relation between scores on type of discipline and criticism of others was also analyzed for each of the groups. Finally, the difference in criticism of others between the positively disciplined subjects from each of the two groups was determined.

TABLE I
RELATIONSHIPS BETWEEN METHOD OF DISCIPLINE AND CRITICISM
OF PARENTS IN THE OVERPROTECTED AND THE
NONOVERPROTECTED GROUPS

		DISCIPLINE		χ^2	p
Group		Negative	Positive		
<i>Overprotected</i>					
Criticism of Parents	(Low ..	3 (7.68)*	11 (6.32)	9.18†	<.01
	(High ..	14 (9.32)	3 (7.68)		
<i>Nonoverprotected</i>					
Criticism of Parents	(Low ..	7 (10.23)	30 (26.77)	4.75†	<.05
	(High ..	6 (2.77)	4 (7.23)		

* Figures in parentheses are expected frequencies.

† Chi square with Yates correction for continuity.

RESULTS AND DISCUSSION

The specificity of the questions and the completeness of the subjects' answers made it possible to score the variables with a high degree of inter-rater reliability. The agreement between the two scorers was: method of discipline (97 per cent), criticism of parents (92 per cent) and overprotected vs. nonoverprotected (95 per cent). The agreement on criticism of others (87 per cent) was not quite as satisfactory. All disagreements were resolved through consensus by conference.

The results in Table I show a significant relation between method of discipline and criticism of parents in both the overprotected ($p = .01$) and nonoverprotected ($p = .05$) groups. This supports the first hypothesis

¹ The writer wishes to acknowledge the assistance of Miss Barbara Schimmel in rating the responses.

CHILD DEVELOPMENT

that subjects whose parents use positive types of discipline have a more favorable attitude toward their parents and family than do subjects disciplined by negative methods. The results are consistent with those reported by Itkin (5), who found, in a study of college students, that for women, the correlation between their attitudes toward their fathers and their attitudes toward the discipline exercised by their fathers was .72; and the correlation between their attitudes toward their mothers and their attitudes toward the discipline exercised by their mothers was .64.

The second hypothesis that subjects whose parents are dominant overprotective and exercise positive methods of discipline displace criticism on to persons other than their parents more than do those subjects disciplined by negative methods was supported. The chi square is significant at the .05 level. The nonsignificant chi square for the nonoverprotected group confirms the third hypothesis of no difference in criticism of others as a function of method of discipline in subjects whose parents are not overprotective.

TABLE 2

RELATIONSHIPS BETWEEN METHOD OF DISCIPLINE AND DISPLACEMENT OF CRITICISM IN THE OVERPROTECTED AND THE NONOVERPROTECTED GROUPS

		DISCIPLINE		χ^2	<i>p</i>
Group		Negative	Positive		
<i>Overprotected</i>					
Criticism of Others	(High ..	5 (8.23)*	10 (6.77)	3.88†	.05
	(Low ..	12 (8.77)	4 (7.23)		
<i>Nonoverprotected</i>					
Criticism of Others	(High ..	5 (4.70)	12 (12.30)	0.02†	—
	(Low ..	8 (8.30)	22 (21.70)		

* Figures in parentheses are expected frequencies.

† Chi square with Yates correction for continuity.

Comparison of the results in Table 2 for the overprotected group to those in Table 1 for the overprotected group indicates that those subjects whose parents use positive methods of discipline criticize their parents less and criticize other people more than do those subjects whose parents use negative methods of discipline. This relationship is interpreted as evidence consistent with a theory of conflict and displacement.

This reasoning is given further support by the results in Table 3 which relate to the fourth hypothesis of the study. Of subjects whose parents exercised positive methods of discipline, those who were overprotected criticized others more than those subjects who were not overprotected. Chi square was significant at the .05 level.

CHARLES Y. NAKAMURA

TABLE 3

RELATIONSHIP BETWEEN TYPE OF CHILD REARING AND DISPLACEMENT
OF CRITICISM IN SUBJECTS WHOSE PARENTS EXERCISE
POSITIVE METHODS OF DISCIPLINE

<i>Group</i>	CRITICISM OF OTHERS		χ^2	<i>p</i>
	<i>High</i>	<i>Low</i>		
Overprotected	10 (6.42)*	4 (7.58)	3.89†	.05
Nonoverprotected	12 (15.58)	22 (18.42)		

* Figures in parentheses are expected frequencies.

† Chi square with Yates correction for continuity.

It is recognized that a possible limitation in this study is the lack of an independent measure of methods of discipline and overprotective or non-overprotective techniques of child rearing practices of the parents of the subjects. However, this fact does not necessarily decrease the confidence which can be placed in the validity of the results. The reliability of children's and college students' responses to questionnaires on home life and parent-child relations is usually high (3, 5). In a study by Meyers (9), the validity of such responses was also found to be quite satisfactory. Meyers developed a questionnaire dealing with several aspects of home life and parent-child relations, including discipline and supervision, which was administered to high school students. Two guidance workers in the school and two social workers in the community who were thoroughly familiar with the homes of 20 of the students indicated their agreement or disagreement with the answers given by the students to the items on the questionnaire. There was complete agreement on 85 per cent of the items between the students and the raters. This provides some assurance for the validity of subjects' reports on aspects of their home environment.

An objective of this study was to determine whether Miller's conflict and displacement theories would be useful in the explanation of certain aspects of parent-child relations which were described by students in the process of receiving counseling for personal problems. The results are encouraging. There is, however, an alternative explanation of the relation between aggressive behavior and methods of discipline. Several writers have offered the concept of identification for explaining children's patterns of behavior (7, 11, 14, 15). Since the identification hypothesis suggests that the pattern of aggression expressed by the parent serves as a model for the child, it is most useful in describing such general patterns as they are observable in the parent. It is less adequate in explaining the reasons for inhibition of aggression and it is especially limited in predicting the direction which aggressive responses may take. It is believed that of the two orientations, the conflict and displacement paradigms have the greater potentiality for fulfilling these latter functions.

CHILD DEVELOPMENT

SUMMARY

The present study tested four hypotheses concerning relationships between direction of expression of hostility by children and methods of discipline and control exercised by their parents. The hypotheses were stated in terms of Miller's conflict and displacement theories as follows:

1. Subjects whose parents exercise positive types of discipline report more favorable attitudes toward their parents than do subjects who are disciplined by negative methods.
2. Of subjects who have dominant overprotective parents, those disciplined by positive methods displace criticism to persons other than their parents more than do those disciplined by negative methods.
3. Of subjects whose parents did not use overprotective techniques of child rearing, there is no difference between those subjects disciplined by positive and those disciplined by negative methods in the amount of criticism displaced toward persons other than their parents.
4. Overprotected and positively disciplined subjects displace more criticism toward persons other than their parents than do nonoverprotected and positively disciplined subjects.

The subjects were 78 college women in an introductory course in personal adjustment. Measures were obtained of parents' method of discipline, parents' overprotective or nonoverprotective techniques of child rearing, children's criticism of parents, and children's criticism of people other than their parents. Scores on these variables were based upon responses to selected questions contained in an autobiographical work book designed for use in college introductory courses in the psychology of adjustment.

It was concluded that the results, which supported all four hypotheses, provided evidence for the usefulness of the conflict and displacement theories in explaining the inhibition of aggressive responses or the direction which they may take.

REFERENCES

1. BANDURA, A., & WALTERS, R. H. Relationship of child rearing antecedents to adolescent behavior disorders. Unpublished paper, Dept. of Psychology, Stanford Univer., 1957.
2. DOLLARD, J., & MILLER, N. E. *Personality and psychotherapy*. New York: McGraw-Hill, 1950.
3. HAYWARD, R. S. Child's report of psychological factors in the family. *Arch. Psychol.*, N. Y., 1935, No. 189, 1-75.
4. HOLLENBERG, ELEANOR, & SPERRY, MARGARET. Some antecedents of aggression and effects of frustration in doll play. *Personality*, 1951, 1, 32-43.
5. ITKIN, W. Relationships between attitudes toward parents and parents' attitudes toward children. *J. genet. Psychol.*, 1955, 86, 339-352.

CHARLES Y. NAKAMURA

6. LEHNER, G. F. J. *Explorations in personal adjustment*. New York: Prentice-Hall, 1949.
7. LEVIN, H., & SEARS, R. R. Identification with parents as a determinant of doll play aggression. *Child Developm.*, 1956, 27, 135-153.
8. LEVIN, H., & TURGEON, VALERIE F. The influence of the mother's presence on children's doll play aggression. *J. abnorm. soc. Psychol.*, 1957, 55, 304-308.
9. MEYERS, T. R. Intrafamily relationships and pupil adjustment. *Teach. Coll. Contr. Educ.*, 1935, No. 651.
10. MILLER, N. E. Theory and experiment relating psychoanalytic displacement to stimulus-response generalization. *J. abnorm. soc. Psychol.*, 1948, 43, 155-178.
11. PAYNE, D. E., & MUSSEN, P. H. Parent-child relations and father identification among adolescent boys. *J. abnorm. soc. Psychol.*, 1956, 52, 358-362.
12. SEARS, R. R. Relation of fantasy aggression to interpersonal aggression. *Child Developm.*, 1950, 21, 5-6.
13. SEARS, R. R., MACCOBY, ELEANOR E., & LEVIN, H. *Patterns of child rearing*. Evanston: Row, Peterson, 1957.
14. STOKE, S. M. An inquiry into the concept of identification. *J. genet. Psychol.*, 1950, 76, 163-189.
15. STOLZ, LOIS MEEK. *Father relations of war-born children*. Stanford: Stanford Univer. Press, 1954.

PARENTAL ATTITUDES AND CHILD ADJUSTMENT

DONALD R. PETERSON,* WESLEY C. BECKER, LEO A. HELLMER,
DONALD J. SHOEMAKER, and HERBERT C. QUAY¹

University of Illinois

Many contemporary personality theories attach great importance to the role parents play in determining the personality characteristics of children. Such emphasis is wisely placed. The primacy, the intimacy, and the extensive protraction of parental influences are likely to render them crucial to the formation of personality tendencies among children. In view of these considerations, there is a clear need for research in which important aspects of parental influence are examined, the behavior of children concurrently appraised, and relationships between the two sets of variables determined.

It is neither possible nor necessary to review here the mountainous literature on parent-child relationships which has emerged over the past 30 years, but certain remarks concerning that literature can and should be made. A good deal of it consists of "expert" advice whose content changes with fashion (13) and whose factual basis is obscure. Much of the rest is theory, which is sometimes a product of keen observation and closely reasoned thought, but which is seldom buttressed by carefully gathered, rigorously evaluated empirical data. There remain a number of excellent investigations on parent-child relationships, among which the works of Radke (10), Baldwin and his collaborators (3, 4), and Sears *et al.* (12) seem particularly outstanding, but it is no disparagement of these and similar studies to note that they leave large areas of the domain almost entirely untouched. A major hiatus exists, for example, in regard to the attitudes of fathers and their part in the formation of personality tendencies among children. A re-

* Department of Psychology, University of Illinois, Urbana.

¹ This project was supported by the Graduate Research Board of the University of Illinois. The writers are grateful to the participating parents and teachers for the cooperation they gave throughout the study, and to Ina Cepenas, Marie Zimmerman, Sally Beck, Herbert Blaufarb, Byron Lindholm, and Dean Froelich for their help in gathering and analyzing the data.

CHILD DEVELOPMENT

view of the literature on parent-child relationships over the years 1929 to 1956 revealed at least 169 publications dealing with relationships between mothers and their children. Available information on father-child relationships, by contrast, was encompassed in 10 articles, one convention address, and one book. This imbalance has ordinarily been justified through reference to an assumption that mothers play a more important part than fathers in the development of child personality. Levy's statement is typical: "It is generally accepted that the most potent of all influences on social behavior is derived from the social experience of the mother" (8, p. 3). The present study permits evaluation of the validity of this assumption through examination of both parents and assessment of their relative influence on the behavior of children.

An analogous sampling restriction obtains in respect to most available investigations of child behavior. Attention has ordinarily been limited either to problem behavior or to "normal" activity. To our knowledge, there has been *no* research in which both fathers and mothers have been studied and their attitudes related to the behavior of disturbed as well as "normal" children, although the methodological advantage to be gained through expansion of variance among children is obvious. In this investigation, the families of both problem and nonproblem children have been examined. The study has two specific purposes. The first is to assess differences between parental attitudes in two groups of families, one in which the children display certain adjustment problems, and another in which they do not. The second is to establish, within the group of families where the children have problems, differential attitudinal patterns associated with two major dimensions of child behavior difficulty.

SUBJECTS AND PROCEDURES

Thirty-one families were selected for investigation from the clientele of a guidance clinic. All the children were from 6 to 12 years of age, of Caucasian extraction, and manifested difficulties in adjustment of sufficient severity to warrant treatment. Cases of known organic brain damage or other serious physical defect, families in which the parents were separated, and cases of intellectual retardation were excluded. We shall refer to this sample henceforth as the *Clinic* group. They were compared with 29 *Non-clinic* families in which the children had been judged by teachers to display acceptable adjustment tendencies in school. Age was again restricted to the range 6 to 12. Since no children with organic disturbances or intellectual defect appeared on the lists submitted by teachers, exclusion in terms of these criteria was unnecessary. Other characteristics of the two samples are given in Table 1.²

² Some of the figures in Table 1 are based on less than the total number of cases, because data were missing. In no instance, however, is $N < 25$.

TABLE 1
SAMPLE CHARACTERISTICS

	CLINIC GROUP		NONCLINIC GROUP	
	Mean	Sigma	Mean	Sigma
Age of child	8.7	2.1	9.1	2.1
IQ of child	107.0	15.5	113.2	14.3
SE status (Warner)	34.7	10.4	35.1	9.6
Age of mother	35.5	6.0	38.6	5.9
IQ of mother	117.7	15.2	115.2	15.3
Education of mother	13.3	2.2	13.8	2.0
Age of father	37.2	5.2	39.9	5.8
IQ of father	124.3	15.0	124.0	15.8
Education of father	14.7	2.9	15.4	3.0
Sex of child				
Number of boys	22		16	
Number of girls	9		13	

The groups do not differ significantly in respect to any of the dimensions examined, though there is an unequal representation of boys and girls in the Clinic group which approaches statistical reliability. The proportion of boys to girls in the latter sample is more than 2 : 1, while there are only a few more boys than girls in the Nonclinic group. Such disparity is typical of most clinic populations, but its possible effect on the results of this study will receive later comment. The discrepancy in child IQ also approaches significance ($.10 < p_t < .20$), but the fact that the IQs of parents were not reliably different (means for Clinic parents were in fact numerically superior to those for the Nonclinic group) suggests that the lower functional intelligence of the Clinic children is part and parcel of the adjustment difficulty, and constitutes an interesting trend in its own right.

On the basis of a one-hour interview, parents were rated on 17 of the 30 Fels Parent Behavior Rating Scales (3). The 17 scales were selected on the basis of Roff's factor analysis (11) in such a way that each factor was represented by at least two of the most heavily saturated rating scales. The factors with which the present analysis is concerned, the elemental variables from which factor scores were computed, the relevant loadings in Roff's analysis, and the weights assigned each variable in derivation of factor scores, are presented in Table 2.

Four staff psychologists and one advanced graduate student in clinical psychology performed the ratings immediately after each interview. Methodological ideals would have been most neatly met by rigorous structuring

CHILD DEVELOPMENT

of the interview, but the need to allow parents of disturbed children to discuss problems in their own terms prevented such standardization. Rather free discussion was therefore permitted through much of the interview, but certain standard questions were asked, as needed, to insure collection of data relevant to the ratings. Distributions of scores on each of the atti-

TABLE 2
INDICANTS OF PARENTAL ATTITUDE

	Loading	Weight
<i>Factor 1—Concern for child</i>		
21. Protectiveness69	3
7. Duration of contact69	2
25. Solicitousness53	1
<i>Factor 2—Democratic guidance</i>		
12. Justification of policy64	2
13. Democracy62	2
24. Readiness of explanation62	2
18. Noncoerciveness of suggestion61	2
<i>Factor 3—Permissiveness</i>		
11. Nonseverity of penalties61	3
10. Nonreadiness of enforcement51	2
<i>Factor 4—Parent-child harmony</i>		
16. Freedom from disciplinary friction63	2
15. Effectiveness of discipline61	2
<i>Factor 5—Sociability-adjustment</i>		
4. Sociability70	3
1. Adjustment55	2
<i>Factor 6—Activeness</i>		
2. Activity76	3
3. Coordination53	2
<i>Factor 7—Readiness of suggestion</i>		
17. Readiness of suggestion58	3
4. Unsociability41	2

tudinal factors were then formed separately for fathers and mothers, and χ^2 analyses of discrepancy between the Clinic and Nonclinic groups conducted. Cutting points were set as close to the first and third quartiles as the distributions would permit, thus segregating the groups into those manifesting high, average, and low scores on each factor, and χ^2 calculated in accordance with the procedure suggested by McNemar (9) for analysis of $2 \times k$ tables.

Information pertinent to reliability of the attitudinal factors was obtained by recording three interviews and comparing judgments made by the five interviewers after listening to the recordings. Over all comparisons, the judges were in agreement more than 80 per cent of the time in classifying subjects as high, average, or low on the factors, and *no* radical disagreements were found, i.e., in no case did a single judge regard a case as high on a given factor when another judge assigned the subject a low score. Such agreement seems quite remarkable in the light of commonly expressed views about the unreliability of clinical judgments. It appears that if raters are asked to judge certain well-defined variables at an appropriate level of differentiation, and are given enough information to do so, they can perform the rating task in a thoroughly acceptable way.

During the interviews with parents, information was also obtained about the nature of the problem which each child displayed. A rating schedule derived from Himmelweit's analysis (7) of Ackerson's data (1) on problem children was employed. The schedule required 3-point ratings of 8 behavior characteristics loading on a *personality problem* factor (sensitivity, absent-mindedness, seclusiveness, day-dreaming, inefficiency in work, inferiority feelings, changeability of mood, and nervousness), and 11 tendencies loading on a *conduct problem* factor (truancy from home, truancy from school, stealing, fighting, lying, destructiveness, swearing, disobedience, rudeness, selfishness, and temper tantrums). Unit weights were assigned, and the two factor scores derived by direct addition of the appropriate elemental scores. The child problem scores were then correlated with each parent attitude factor to determine the kinds of attitudes associated with each problem dimension. This analysis was performed separately for mothers and fathers, and was restricted to the Clinic group. To reduce contamination, attitudes of mothers were correlated with judgments of child behavior based on the reports of fathers, and attitudes of fathers correlated with judgments based on the mothers' reports of child behavior.

Reliability of ratings on problem activity was estimated by correlating scores derived from the interviews of mothers with those based on the accounts of fathers. Correlations of .40 and .83 emerged for personality problem and conduct problem, respectively. Evidently the open behavior involved in conduct problems can be judged with fair accuracy, but the fine discriminations required in judging the severity of internalized difficulties, within such a homogeneous group as the present one, are too difficult to permit close agreement.

It is of methodological interest to note that gross classification of children into those manifesting personality problems and those manifesting conduct problems could be done in a highly reliable way. When classification was made simply by noting which of the two problem scores was higher, and designation based on father report compared with that based on mother report, only one case of actual disagreement was found. Classification was impossible in four cases, because the scores were equal, but in-

CHILD DEVELOPMENT

dependent judgments were in accord for the remaining 26 cases (84 per cent). Insufficiency of *N* in the two groups of problem children prohibited direct comparative study, and required use of the correlational analysis of dimensions described above, but investigators can be encouraged by the possibility of attaining respectably high agreement, not only in rating parental attitudes, but in rating child behavior, when variables are defined with proper care and the differentiations required are appropriate to the sensitivities of the judges.

Configural analyses of inter- and intraparental patterns were considered, but on the basis of two pilot studies (5, 14) rejected in favor of the simpler,

TABLE 3
COMPARISON BETWEEN CLINIC AND NONCLINIC PARENTS

Parent Attitude Factor	M O T H E R S				F A T H E R S			
	Clinic	Non-clinic	χ^2	<i>p</i>	Clinic	Non-clinic	χ^2	<i>p</i>
1. Concern for child			1.92	<.50			5.40	<.10
high	10	5			12	6		
average ..	15	16			9	17		
low	6	8			10	6		
2. Democratic guidance			8.58*	<.02			10.25*	<.01
high	5	11			4	13		
average ..	13	15			15	13		
low	13	3			12	3		
3. Permissiveness			4.79	<.10			1.17	<.70
high	8	6			8	9		
average ..	11	18			14	15		
low	12	5			9	5		
4. Parent-child harmony			21.13*	<.001			19.03*	<.001
high	2	11			4	13		
average ..	15	18			13	16		
low	14	0			14	0		
5. Sociability-adjustment			18.09*	<.001			7.14*	<.05
high	2	14			5	10		
average ..	15	13			14	16		
low	14	2			12	3		
6. Activeness			1.04	<.70			8.21*	<.02
high	10	6			10	4		
average ..	15	16			11	21		
low	6	7			10	4		
7. Readiness of suggestion			2.20	<.30			7.69*	<.05
high	10	5			13	3		
average ..	15	15			13	18		
low	6	9			5	8		

* Significant at or beyond .05 level.

PETERSON, BECKER, HELLMER, SHOEMAKER, and QUAY

more straightforward, and evidently more powerful statistical techniques described above. A factor analysis drawn from the same basic data pool as the present one but dealing with a differently constituted set of variables is reported elsewhere (4):

RESULTS

Results of comparison between the Clinic and Nonclinic parents are given in Table 3. Both mothers and fathers of problem children were judged to be less well adjusted and sociable, more autocratic, and to experience more disciplinary contention than were the parents in the Nonclinic group. No other reliable discrepancies were found for mothers, although a tendency for Clinic mothers to be either very strict or very permissive in disciplining their children approached significance. Fathers, however, differed reliably in respect to two dimensions which did not differentiate between the Clinic and Nonclinic mothers. The fathers of problem children were more prone to make suggestions than were the fathers of children without known problems, and were either highly active and rigidly organized or relatively inactive and disorganized in the conduct of their affairs. They also tended toward extremes in regard to concern for their children, but discrepancy along this dimension did not quite reach the level usually demanded for assertion of statistical significance.

TABLE 4
CORRELATIONS BETWEEN PARENT ATTITUDE FACTORS AND
CHILD PROBLEM DIMENSIONS

Parent Attitude Factor	MOTHERS		FATHERS	
	Personality Problem	Conduct Problem	Personality Problem	Conduct Problem
1. Concern for child	-.19	.12	-.21	.28
2. Democratic guidance	.03	.06	-.34	.28
3. Permissiveness	-.12	.22	-.14	.21
4. Parent-child harmony	.33	-.58*	.41*	-.70*
5. Sociability-adjustment	.19	-.41*	-.09	.05
6. Activeness	-.03	.19	.36*	-.09
7. Readiness of suggestion	.04	-.04	.06	-.02

* Significant at or beyond .05 level.

Correlations between each of the two dimensions of child disturbance and each parent attitude are presented in Table 4. Two considerations must be kept in mind in interpreting these data. The first is that the correlational index used (product-moment r) requires an assumption of linearity which was imperfectly met in regard to certain of the variables. Such fail-

CHILD DEVELOPMENT

ures may have obscured a few actually significant relationships, but inspection of scatter plots revealed no obvious curvilinearities, and use of r was considered appropriate. The second consideration is that all correlations are derived from the Clinic group only. Expansion of variance through including Nonclinic families would almost surely have increased some of the correlations to a significant level. It seemed appropriate to restrict this particular analysis to Clinic families because it is in respect to such groups that clinicians typically function. Differential diagnosis, prediction, reconstruction of possible causal relationships, and choice of treatment demand discrimination between individual cases *all* of whom present problems of one kind or another, not merely between people who have problems and those who do not.

No matter how they are construed, however, the correlations in Table 4 are of lower magnitude than most previous research and theory would lead one to expect. Three of the five significant figures relate to parent-child harmony, which is at least as much a function of child behavior as of parental attitude, and one of the others is related to parental adjustment, a dimension along which rater bias seems especially likely to operate. Evidently many factors other than parental attitudes, at least as we have measured them, enter into determination of these two problem tendencies in child behavior. Within the family, however, it is again apparent that the paternal role is fully as important as the maternal one, and, if the figures given above are reproducible, fathers may play a slightly more crucial part than mothers in determining not only whether children have problems or not, but the kinds of problems they are likely to develop. The average correlation over all attitudes and both problems is .24 for fathers and .19 for mothers.

In erecting hypotheses for study in future research, it is often fruitful to consider not only statistically significant results, but more modest data trends as well. If this is done by considering all parent attitudes for which r exceeds .20, the following patterns emerge. Personality problems seem largely independent of maternal attitude, but related to dictatorial attitudes and a lack of genuine concern among fathers. Conduct problems are related chiefly to maladjustment among mothers, and to democratic attitudes and heightened feelings of parental concern among fathers. Both parents appear overly permissive, and characterization of the fathers of children who present conduct problems as generally weak and ineffectual may not be far wide of the mark. Democratic qualities are esteemed in our culture, but when they are combined with laxity, unwillingness to issue orders, exaggerated concern for children, and a tendency to shelter them in the face of day-to-day problems³ the seeds of conduct disorders may be sown.

³ Correlations above .20 obtained between conduct problem score and ratings of fathers on the following Fels variables: 10, laxity; 13, democracy; 15, unsuccessful policy; 16, disciplinary contention; 18, noncoerciveness of suggestion; and 21, general protectiveness. The complete matrix of correlations between child problem scores and scores on the elemental Fels variables are not reported here, but can be obtained by writing to the authors.

DISCUSSION

Probably the most significant finding to emerge from this study is that the attitudes of fathers are at least as intimately related as maternal attitudes to the occurrence and form of behavior problems in children. The popular choice of mothers as a focus of research attention and the general clinical tendency to offer psychotherapy to the mothers rather than the fathers of disturbed children are usually rationalized by noting that mothers generally spend more time with their children than fathers do, and by speculating that mother-child relationships are more intimate affairs than those between fathers and children. Whatever the validity of these assumptions, only the final effect is of fundamental concern, and we now have reason to believe that the emphasis in this effect is not as one-sided as it has previously appeared.

The extent to which the present results were affected by sex disparity among children is difficult to assess. Surely the presence of twice as many boys as girls in the Clinic sample could have a bearing on any findings related to relative influence of the two parents on child behavior, and we may only have demonstrated that fathers play a more vital role in influencing the behavior of boys than has previously been assumed. Even this, however, would be a result worth noting, and, while there is an obvious need to find out whether the present results can be repeated in a study involving more nearly equal representation of boys and girls, the necessity for examining paternal influence remains as vital as before. The practical difficulties involved in securing the cooperation of fathers for a study such as this are considerable, but they can be overcome, and it is now clear that they must be overcome if the social environment, in its relationship to personality development in children, is ever to be understood.

The various relationships, as we have reported them, are probably better estimates of "true" association than some theory and previous research would lead one to expect. We have eliminated at least a little of the contamination-generated spuriousness of many earlier investigations, and as a consequence, relationships between parental attitudes and child behavior appear with generally lower magnitude. We would be the last to deny the importance of the relationships we have presented, no matter how small they seem, and in evaluating the importance of our results feel justified in agreeing with Sears *et al.* (12), who regard the location of any real (i.e., replicable) influence, however low its magnitude, as a contribution to knowledge. "If our general assumption is correct—i.e., that any given behavior is the product of many influences—it would be quite impossible to obtain high correlations between single child-rearing dimensions and the measures of child behavior" (12, p. 456). But the results pointedly indicate the operation of many factors other than the ones we chose to examine. Though certain general statistical tendencies emerged, we still found families in which the parents appeared maladjusted, evidently didn't get along, and exhibited the most abhorrent kinds of attitudes toward their

CHILD DEVELOPMENT

children, but the children appeared to be getting along beautifully. We saw parents whose attitudes and other characteristics were in nearly perfect congruence with the stereotype of the "good parent," but whose children displayed problems of the most severe order. The need to expand the scope of studies on personality development, and to examine other parental characteristics, relationships with such influential adults as teachers, interactions with siblings and peers, in sum, a more complete matrix of social factors, is patently clear.

It is our belief, however, that even if *all* social influences could be encompassed, and absolutely perfect measurement of them attained, a sizeable share of the variance in child behavior would still be unexplained. Parents frequently reported that their children had displayed certain behavior consistencies very early in life, and that these tendencies had remained stable in the face of what often appeared to be very extensive changes in social environment. In a study of parent-child relationships of the present kind, we are not dealing with unidirectional causalities; we are dealing with interactions. We have reported a number of correlations and other relationships between parental attitudes and child behavior, and have sometimes implied that the former caused the latter. This inference seems the most likely of the several which might be made, but the tenability of other interpretations cannot be denied. To explicate the second most likely inference, it is often just as reasonable to assume that personality tendencies on the part of the children, appearing very early in life and possibly of constitutional origin, have engendered modification of parental attitudes. The parents of a stable, predictable, sensible child can afford to be democratic. As Escalona (6) has suggested, the parents of an erratic, difficult, peculiar child may become apparently inconsistent out of sheer desperation. The direction of a causal relationship can only be established through determination of temporal priority, and the need to examine children *very* early in their development, and to explore the effects children may have on parents, is logically as critical as examination of the kinds of relationships with which this study is primarily concerned.

The extent to which the results have been influenced by rater bias is unknown. The interviewers made a conscious effort to avoid such bias, and it is entirely possible that estimates of "true" relationships were reduced rather than magnified through the operation of the systematic error such an effort could conceivably entail. In the correlational analysis, contamination was reduced by relating parental attitudes as measured on interview with one parent to child behavior as independently assessed by the clinician who saw the other parent. Throughout the study, however, interviewers knew which parents belonged to which group, they had rather firm ideas about the kind of behavior disturbance each child in the Clinic group presented, and there is a real possibility that some of the "significant" results reported above are spurious. Certain of the results were contrary to expectation, no obvious "bias factor" emerged on factor analysis of these and other data (4), and both considerations suggest that obtained associa-

tions were not completely generated by the preconceptions of the raters. Still the actual effect, even the probable direction of the effect of bias, is indeterminate, and will remain so until the appropriate methodological improvements are made. The need for direct, independent, objective measurement of parental attitudes and of child behavior is obvious. Much of our current effort is addressed to the development of such measures in the conviction that their proper application can determine some of the facts so urgently needed for adequate explanation, accurate prediction, and successful modification of developing personality trends in children.

SUMMARY

This study was designed to furnish information about relationships between parental attitudes and child behavior. It began with two purposes. The first was to assess differences between parent attitudes in two groups of families, one in which the children displayed certain adjustment problems (Clinic group, $N = 31$), and another in which no such problems had been observed (Nonclinic group, $N = 29$). The second aim was to establish differential attitudinal patterns associated with two major dimensions of child behavior disturbance. On the basis of interviews with mothers and fathers, ratings pertinent to seven parent attitude factors were made. Attitudes in the Clinic group were then compared with those in the Nonclinic group, separately for mothers and for fathers, and correlations were computed between the measures of parent attitude and indirect but independent evaluations of child problem behavior. The following results emerged:

1. Contrary to general assumption and our own original expectation, the attitudes of fathers were found to be at least as intimately related as the attitudes of mothers to the occurrence and form of maladjustive tendencies among children.
2. Both mothers and fathers of children who displayed adjustment difficulties were judged to be less well adjusted and sociable, less democratic, and to experience more disciplinary contention than the parents of children with no manifest problems.
3. In addition, Clinic fathers were regarded as more prone to offer suggestions, and tended toward extremes along a dimension of activity and organization in the conduct of their affairs.
4. Personality problems among children in the Clinic group were found to be relatively independent of maternal attitudes, but appeared to be related to autocratic attitudes and lack of parental concern among fathers.
5. Conduct problems were associated with general maladjustment among mothers in the Clinic group, and with evident permissiveness and disciplinary ineffectuality on the part of fathers.

Some rather impressive research has been done on mother-child relationships. But there is an equally urgent need to examine other social factors, as well as constitutional predispositions, in personality development,

CHILD DEVELOPMENT

ideally through the application of objective measures. The need to study fathers, as crucial components of the social-environmental matrix from which child personality tendencies emerge, seems especially vital.

REFERENCES

1. ACKERSON, L. *Children's behavior problems*. Chicago: Univer. Chicago Press, 1942.
2. BALDWIN, A. L. The effect of home environment on nursery school behavior. *Child Develpm*, 1949, 20, 49-62.
3. BALDWIN, A. L., KALHORN, JOAN, & BREESE, FAY H. The appraisal of parent behavior. *Psychol. Monogr.*, 1949, 63, Whole No. 4.
4. BECKER, W. C., PETERSON, D. R., HELLMER, L. A., SHOEMAKER, D. J., & QUAY, H. C. Factors in parental behavior and personality as related to problem behavior in children. *J. consult. Psychol.*, in press.
5. CEPENAS, INA. Maternal attitudes and child behavior. Unpublished B.A. thesis, Univer. of Illinois, 1957.
6. ESCALONA, SIBYLLE. Some considerations regarding psychotherapy with psychotic children. *Bull. Menninger Clinic*, 1948, 12, 127-134.
7. HIMMELWEIT, HILDE T. A factorial study of "Children's behavior problems." Unpublished manuscript, Univer. of London, 1952. Cited in Eysenck, H. J. *The structure of human personality*. London: Methuen, 1953.
8. LEVY, D. M. *Maternal overprotection*. New York: Columbia Univer. Press, 1943.
9. MCNEMAR, Q. *Psychological statistics*. New York: Wiley, 1949.
10. RADKE, MARION J. *The relation of parental authority to children's behavior and attitudes*. Minneapolis, Minn.: Univer. of Minnesota Press, 1946.
11. ROFF, M. A factorial study of the Fels Parent Behavior Scales. *Child Develpm.*, 1949, 20, 29-45.
12. SEARS, R. R., MACCOBY, ELEANOR E., & LEVIN, H. *Patterns of child-rearing*. Evanston, Ill.: Row, Peterson, 1957.
13. STENDLER, CELIA B. Sixty years of child training practices. *J. Pediat.*, 1950, 36, 122-134.
14. ZIMMERMAN, MARIE. Paternal attitudes and child behavior. Unpublished M.A. thesis, Univer. of Illinois, 1958.

PERMISSIVENESS, PERMISSION, AND AGGRESSION:
THE EFFECT OF ADULT PRESENCE OR ABSENCE
ON AGGRESSION IN CHILDREN'S PLAY

ALBERTA ENGVALL SIEGEL*¹

Pennsylvania State University

and LYNETTE GAYLE KOHN¹

Stanford University

It has become conventional to think of an adult's permissiveness toward a child as representing neutralization of adult control. By behaving permissively (acceptantly, nonjudgmentally, etc.), the adult is thought to be "freeing" the child from his presence and authority, and in some loose sense to be permitting the child to behave as he would in the absence of any adult. Such thinking has emphasized the punitive or threatening aspects of adult-child relations. Adult permissiveness is seen as a way of reducing the child's fear of adult punishment and thus as a way of encouraging him to express in behavior those impulses whose expression he knows to be normally unacceptable to adults.

Such thinking about permissiveness has been the basis for explanations which have been advanced for a frequent finding about aggression in doll play. It has repeatedly been observed that aggression in doll play increases over time or from session to session (1, 2, 5, 6, 8, 9). This finding has typically been explained by reference to the adult experimenter's acceptant and nonjudgmental interaction with the child and the fear-reducing effect this interaction has. For example, in discussing recent evidence confirming the finding, Levin and Sears state that

For both sexes, there was a significant session-to-session increase in the proportion of aggression units. This finding . . . has been interpreted to mean that the deliberate permissiveness of the experimental procedure progressively acts to reduce the inhibitions about aggression that the child has heretofore acquired (2, pp. 149-150).

*Department of Child Development and Family Relationships, Pennsylvania State University, University Park.

¹ We acknowledge with gratitude the skilled services of Miss Lyn Kuckenberg as the observer in this study, and the generous cooperation of Dr. Edith M. Dowley, Director, and Miss Patricia Rowe, Head Teacher, of the Stanford Village Nursery School, where the study was conducted.

CHILD DEVELOPMENT

Support for the notion that the increase of aggression over time "might indicate a lessening of inhibition on the part of the child, due to the fact that his earlier aggressive actions have met with no criticism from the experimenter" (6, p. 164) is given by two findings from Pintler's study. The first is that the latency of aggressive acts is significantly shorter under a condition of high experimenter-child interaction (which was deliberately permissive and nonjudgmental) than under a condition of low experimenter-child interaction. The second is that the amount of thematic aggression is significantly greater under conditions of high experimenter-child interaction.

If the increase over time is explainable in terms of diminution of fear alone, we would expect an increase of aggression from session to session in the *absence* of an adult as well as in an adult's presence. This expectation, however, was disconfirmed in a study of children's social play in two sessions in the absence of any adult (12). Contrary to the prediction that would be drawn from the assertion that inhibition of aggression is caused by fear of adult punishment, it was observed that aggression *decreased* from session to session in the absence of any adult. In the light of the findings already mentioned, this finding was interpreted to mean that whether aggression increases or decreases with time depends on whether a permissive adult is present, for

The evidence of this study . . . suggests that the presence of a permissive adult in a play session may have a cumulatively facilitating or releasing effect on children's aggression. . . . This interpretation . . . forces a somewhat more positive conception of the result of experimenter permissiveness than might otherwise be held. Lack of fear of punishment is not a sufficient explanation. . . . A more adequate explanation seems to be that in the presence of an adult experimenter the young child abdicates superego and ego control functions to him, whereas in the absence of any adult the child's own internalized standards are invoked (12, p. 376).

The explanation of the finding by reference to adult presence or absence, however, was admittedly based on only rough evidence, for it was drawn from a comparison of session difference findings in studies which differed in other ways than in the presence or absence of an adult. The increase in aggression from session to session was observed in doll play studies of a single child in the presence of an adult; the decrease in aggression from session to session was observed in a study of the social play of two children in the absence of an adult. This means that the interpretation of these findings together as demonstrating the positive or facilitating effect of adult permissiveness "may properly be viewed as an hypothesis which should be tested in a direct comparison of two play situations which differ only in the presence or absence of a permissive adult" (12, p. 377). The present paper reports a study designed to permit just such a direct comparison.

Rather than thinking of permissiveness as adult self-neutralization, we may conceive of permissiveness positively as a way of expressing judgments

or employing imputed authority. According to this conception, adult behavior which is intended by the adult to be "acceptant" is seen by the child as permission-giving, and an adult attempt at being "nonjudgmental" is interpreted by the child as being in fact affirmatively judgmental.

A conception of permissiveness approaching this one is used by Sears, Maccoby, and Levin (10) in their ratings of mothers' statements about their practices concerning aggression. A mother is categorized as "entirely permissive" if she "never interferes, never tells child she does not want him to fight," and if she considers aggression to be "natural, part of growing up" (10, p. 243). These researchers found that, according to maternal reports, high permissiveness in the sense just defined is associated ($r = .23$) with high aggression in children. They account for this association by referring to the mutual expectations of mother and child as well as to punishment and fear of punishment:

When a mother adopts a permissive point of view about aggression, she is saying to her child, in effect, "Go ahead and express your angry emotions; don't worry about me." She gives few signals in advance that would lead the child to fear to be aggressive. On the contrary, her behavior is one of expectancy that he *will* be, and that such behavior is acceptable. It is scarcely surprising that a child tends to fulfill her expectations (10, p. 259).

An adequate account of aggressive behavior must acknowledge the force of a number of influences in determining the amount and kind of aggression a child exhibits. One of these, of course, is the strength of his aggressive drive, and this is presumably a resultant of a range of experiences in his personal history and of the operation of certain characteristics of the immediate setting.

A second influence of importance is the child's own attitude toward aggression—his own evaluation of its goodness or badness. This may be presumed to be a distillation and internalization of the attitudes of significant people in his life. These two influences—the strength of the child's aggressive drive and the nature of his attitudes toward aggressiveness—together undoubtedly account for much of the variability in aggression from child to child and also for at least some of the systematic variability between the sexes.

A third influence is the expressed attitudes and values of the other people in the setting in which aggression may be displayed. In our culture, with its extremely complex code about when, where, how, and to whom aggression may appropriately be displayed, people learn to derive cues about what behavior is acceptable at any given time and place by observing the behavior and communications of others in the setting. Children learn especially to rely on adult behaviors and communications to provide hints as to what is suitable or appropriate. This third sort of influence probably accounts for many differences in children's behavior from one setting to another, and also is probably important in shaping sex differences in both incidence and modes of aggression.

CHILD DEVELOPMENT

A fourth influence is the expectation that aggression will elicit counter-aggression (punishment). Presumably, such an expectation typically acts to inhibit overt expression of aggression.

It is by reference to the fourth factor that the finding of an increase in aggression from session to session has been explained. We believe that an adequate account of both the session-to-session increase in an adult's absence must include reference to all of the influences which have been enumerated, and especially the second, third, and fourth.

In either the presence or the absence of an adult, we may expect wide individual differences in incidence of aggression, because of the wide range of strength of aggressive drives in children and because of individual differences in attitudes toward aggression. In addition to these individual differences irrespective of condition, however, we may expect systematic group differences dependent on the condition of adult presence or absence.

In the presence of a permissive adult, a child's aggression may be expected to increase because (a) he will infer from the behavior of the adult that this is a setting in which aggression is suitable or appropriate, and (b) with experience he will undergo a progressive decrease in inhibition based on fear of punishment. That is, changes in the nature of the third and fourth influences mentioned above account for an increase over time.

In the absence of any adult, however, a child's aggression may be expected to decrease from session to session. Lacking any adults to define the social situation and to express expectations of his behavior in it, the child will, after an initial "release" or testing of the limits, rely increasingly on his own learned standards of conduct (the second influence mentioned above) which, in the middleclass child, will typically be that aggression is unacceptable in social play.

HYPOTHESIS

The study reported here was conducted to test the hypothesis derived from the above account, one concerning session-to-session differences in aggression in children's play in the absence of an adult relative to such differences in the presence of an adult. The hypothesis is that *children under the two conditions will exhibit different session-to-session changes in aggression, in that the aggression of the children under the adult-absent condition will tend to decrease in comparison to the aggression of the children under the adult-present condition, which will tend to increase.*

METHOD

Subjects

The Ss were boys enrolled in a university nursery school. Boys were used exclusively because the amount of aggression they display in the play setting used in the study is much greater than the amount displayed by girls (11,

p. 374) and therefore boys' aggression scores may be presumed to be more sensitive to change. The boys served in the experiment in pairs, both members of any pair being from the same play group in the school. One member of each pair was chosen from the older boys in the group (among the upper 50 per cent in age) and one from among the younger. The boys were paired by the Director of the nursery school on the basis of their usual friendliness to each other and the supportiveness they could be expected to have for each other in an unfamiliar play setting. The older boy of each pair was a *S* in the experiment. Eighteen pairs were selected, and the older boy: (*Ss*) in these pairs ranged in age from 4-1 (4 years, 1 month) to 5-0. The younger boys (who were not *Ss*—their aggression in play was not scored) ranged in age from 2-11 to 4-9, the great majority being 3-year-olds or early 4-year-olds. There were three independent play groups in the school; therefore, some minor overlap occurred between the age ranges of "older" and "younger" boys.

After the pairs had been selected, they were assigned at random (by the toss of a coin) to the two conditions: Adult-Absence and Adult-Presence. Nine pairs were assigned to each condition. The average age of *Ss* in both groups was 4-7.

Procedures with Subjects

Each pair participated in two play sessions. These occurred two days apart. Any pair's two sessions were held at the same time of day, insofar as this could be accomplished without destroying the informal and permissive nature of the *E*'s interactions with the *Ss*. (In the Adult-Absent group, the difference between time-of-day for initiation of one session and time-of-day for initiation of the other was 22 minutes on the average, and these differences ranged from 0 to 50 minutes. In the Adult-Present group, the average difference in time-of-day of initiation was 16 minutes, with a range from 0 to 40 minutes.)

All play sessions were conducted in a small playroom in the same building as the nursery school classrooms. This room is reserved for testing and research sessions, and is familiar to the children in the school.

The sessions were initiated in an identical manner for all pairs. *E* invited the pair to come to the playroom so that she could read a story to them and then let them play with some special toys there. As soon as the children entered the room, they were asked to be seated in two chairs, and *E* read a story to them.

The two stories used for the two sessions were rather similar in content, each being a simple animal story which was neutral with respect to aggression and which was printed in a brightly illustrated book. These stories were selected during a pilot study for their interest and appeal for young boys. A counterbalanced design determined which story was read first and which second for any pair.

After the story had been read, *E* invited the children to play as they wished with any of the toys in the room. The toys, selected to be similar

CHILD DEVELOPMENT

to those used in the earlier study (12), were these: two rubber daggers, two small plastic toy guns, one singing spinning top, two lumps of clay, a small wooden train composed of detachable cars, three small metal automobiles, five or six inflated balloons, a small clown toy which was weighted to remain erect, and a large inflated plastic punching toy (a clown) which stood child height. The arrangement of the room, which was furnished with three chairs and a table, was the same at the beginning of each session.

For the pairs in the Adult-Present condition, *E* sat quietly in a chair at one side of the playroom during both sessions. Her attitude was friendly, acceptant, interested, but nonintrusive. She did not initiate any conversations, and she replied only briefly and noncommittally to any initiations from a child.

For the pairs in the Adult-Absent condition, *E* followed her invitation to the children to play with another statement indicating that she had to leave the room for a while to work at her desk. (*E* had shown this desk to the Adult-Absent pairs while walking with them to the playroom; the desk was in a hall some distance from the room.) She explained that they were to stay in the room until she returned for them. She said that she would knock at the door before entering the room upon her return, and that during her absence they would be alone and their privacy would not be disturbed. The purpose of these remarks, which were restated or amplified when necessary to assure their being understood, was to assure the children of freedom from adult intrusion, supervision, or observation during their play.

Some children under the Adult-Absent condition left the room during the play session, usually to carry some item of information to *E* at her desk. In such instances, *E* walked back to the playroom door with them, repeated her instructions, and returned them to the playroom.

Each pair remained in the room for a 14-minute play period. During this period their play was observed by an observer behind the one-way vision mirror on one wall of the room. At the conclusion of the period, *E* terminated the play session and walked with the children back to the nursery school. For the Adult-Absent pairs, her conduct was exactly as she had told them it would be: she returned to the door of the playroom, knocked, and waited for their invitation before entering. Moreover, she ignored the state of the room, and she responded in an uninquisitive and noncommittal way to their remarks about what went on during her absence.

Scoring of Aggression

All play sessions were observed and scored by the same person, a graduate student in psychology who was kept ignorant of the purposes of the study and of the hypothesis under test. The scoring procedure was identical to that reported earlier (11, 12). Scoring began as soon as the instructions to the children were concluded in the Adult-Present sessions, and as soon as *E* left the room in the Adult-Absent sessions. Paced by an electric timer and using prepared scoring sheets, *O* entered a rating of aggression for

every 20-second interval. The rating was either a 0 (no aggression), 1 (mild or playful aggression), 2 (stronger or more forceful aggression), or 3 (intense aggression in which the child seemed highly involved). *O*'s rating of intensity represented, by convention, the *most extreme* aggressive behavior of the child during the 20-second interval. In judging intensity, *O* considered both the quality of the instrumental act and the nature of the goal response; an act might be judged intensely aggressive because its aim was highly destructive or hostile or because its mode of execution was highly forceful and showed much self-involvement. Only the behavior of the older member of each pair, the *S*, was scored. A *S*'s aggression score for a play session was the sum of the ratings he received for the 42 20-second intervals during that session. Thus, theoretically, scores could range from 0 to 126. In fact, the observed range of scores for these *S*s was from 10 to 114. In previous work (11, 12), this scoring method has been shown to have high interobserver reliability: the agreement between total scores from 48 protocols of two independent observers was indicated by the Pearson correlation, $r = .97$. For the present study, pilot sessions were conducted to train *O* in the scoring technique and to conduct a preliminary check on interobserver reliability, which was found to be satisfactory. In the course of the experiment itself, 12 play sessions (including some of each type—first Adult-Absent session, second Adult-Absent session, first Adult-Present session, and second Adult-Present session) were observed and scored independently by a second *O*, and the interobserver agreement for these sessions is given by the Pearson correlation, $r = .98$.

Summary of Experimental Controls

Summarized below are the controls which were instituted in the experiment in order to enable isolation of the effects of adult-presence or adult-absence on session differences in aggression:

1. Any *S*'s partner in play was the same child for both sessions. Thus, any differences between sessions cannot be attributed to systematic differences in interpersonal stimulation.
2. Only one member of each pair served as a *S*. The statistical test's requirement that the scores be independent was thus met.
3. A counterbalanced design determined which story was read first and which second for any pair, so that any possible systematic differential effect of any story would not appear as an experimental effect.
4. The pairs were assigned to the two conditions by a random procedure. Any significant differences between the *S*'s behavior under the two conditions, therefore, cannot reasonably be attributed to any probable prior systematic differences between the *S*s in the two groups.
5. Possible time-of-day effects on incidence of aggression were held constant; any pair participated in both sessions I and II at the same time of day.
6. Timing effects were controlled by separating sessions I and II by a constant interval of time (48 hours) for every pair.

CHILD DEVELOPMENT

7. Possible observer bias was controlled by keeping *O* ignorant of the hypothesis under test. *O* knew whether or not *E* was present at any play session, of course, but she did not know that *E*'s presence or absence was the experimental effect under study, and she was not familiar with the research literature concerning the influence of an adult on children's play.

8. The *E* was the same person for all sessions, and she attempted to interact with all children similarly except with respect to those acts of hers which necessarily depended on whether she was present or absent during the play sessions.

RESULTS

The hypothesis was tested by a Mann-Whitney test (13, pp. 116-127) on the difference scores (each *S*'s session II score for aggression subtracted from his session I score). The data are shown in Table 1. The scores for the *S*s in the adult-absent sessions show changes significantly different ($p = .01$) from those shown by the scores for the *S*s in the adult-present sessions. As inspection of Table 1 will reveal, all *S*s in the adult-absent settings showed *less* aggression in session II than in session I, whereas two-thirds of the *S*s in the adult-present settings showed *more* aggression in session II than in session I. Thus, the hypothesis is confirmed.

TABLE 1
SESSION DIFFERENCES IN AGGRESSION

Subject and Condition	AGGRESSION SCORE		Difference
	Session I	Session II	
Ab-1	65	10	55
Ab-2	73	46	27
Ab-3	58	52	6
Ab-4	90	30	60
Ab-5	56	50	6
Ab-6	80	74	6
Ab-7	103	97	6
Ab-8	41	26	15
Ab-9	97	72	25
Pr-1	46	48	- 2
Pr-2	17	22	- 5
Pr-3	97	114	-17
Pr-4	75	59	16
Pr-5	30	53	-23
Pr-6	55	42	13
Pr-7	86	107	-21
Pr-8	58	39	19
Pr-9	72	93	-21

ALBERTA ENGVALL SIEGEL and LYNETTE GAYLE KOHN

This finding does not imply an over-all difference in incidence of aggression under the two conditions: when each *S*'s scores for the two sessions are considered together, the mean of these totals for the Adult-Absent *Ss* (124.4) is less than a point different from the comparable mean for the Adult-Present *Ss* (123.7). Thus, it is not the total incidence of aggression which was shown to vary under the two conditions, but rather the timing of its occurrence.

DISCUSSION

The results of this study confirm the hypothesis drawn from the contention that adult permissiveness must be conceived in more positive terms than simply as a way of reducing a *S*'s fear of punishment.

In their discussion of ego functions, Redl and Wineman have described a phenomenon similar to that under study. They describe the "spontaneous establishment of substitute controls":

Even in normal children, the control system of the ego does not always have to stay switched on to its full volume just to keep things from getting disorganized. Much of the time, ego vigilance and ego control can be switched back to low, just because there are adequate outside control forces at work. In those cases, the ego gets its flow of support from the presence of authority figures, the soothing awareness of a relaxed and friendly atmosphere, the perception of existing routines or well-oiled rules and regulations. . . . Sometimes, especially when impulsivity runs high, even well-adjusted children have trouble keeping to the level on which they were performing when such "outside controls" suddenly drop out. Thus . . . the teacher leaving the room may find noise rising in the classroom in spite of the warnings or pleas she left behind; the change over from a more highly pressured classroom to one with a wider range of permissiveness may cost ten minutes of temporary disorganization.

The normal child is supposed to have some reserves to institute inside controls quickly after the outside ones have petered out. . . . In fact, we could think of no better test for the emergency vigilance of ego functions than just such moments of withdrawal or breakdown of outside structures or controls (7, pp. 110-111).

In terms of these authors' discussion, *Ss* under the Adult-Absent condition in the present study were observed spontaneously to establish "substitute controls," to "switch to high," to demonstrate the "emergency vigilance of ego functions." On the other hand, *Ss* under the Adult-Present condition could "switch to low," and could get a flow of support from the existence of an accepting authority figure and the perception of rules and regulations consonant with their behavior. It may be suggested that a technique for assessing a child's ego strength or maturity of controls would be to compare his behavior in the two settings of this study (adult-absence and adult-presence) and especially to observe the swiftness and ease with which he "switches to high."

Also relevant here is the psychoanalytic account of the relations between leader and follower. As presented by Munroe, this theory sees a leader as

CHILD DEVELOPMENT

a substitute for internalized controls. "The leader and the admonitions of the leader are substituted for the ego ideal or superego. The great man . . . takes over the individual conscience" (4, p. 146). In the terms of the psychoanalytic account, the present study concerns children who abdicate superego functions to an adult leader when in her presence—"Interestingly enough, people usually do not feel acute remorse for actions committed under conditions of leadership conscience" (4, p. 146)—but who increasingly rely on their own ego ideal or superego for standards of behavior in the absence of any adult.

The importance of the presence of an adult, and the differential effects of adults in different role-relations to the Ss, has also been highlighted by the findings of Levin and Turgeon (3) who studied changes in children's aggression in two sessions of doll play. The first session was attended by only the child and the experimenter. During the second session, the child's mother was in attendance for one group of Ss and an adult woman who was a stranger to the children was in attendance for the other group. Each of the children watched by their mothers was more aggressive in the second than in the first session, whereas four-fifths of the children watched by a stranger were less aggressive. These session differences are in the opposite direction from what had been predicted in a hypothesis drawn from displacement theory, and it is of interest to note that Levin and Turgeon refer to the nature of superego development in the child in their attempt to account for the unexpected finding. They suggest the plausibility of a "transfer of superego" explanation, one which posits that the child returns impulse control to the mother in her presence. "Self control is a worrisome burden for the child, so that he is ready to transfer his newly acquired control to his mother when she is available. Anecdotally, we know that mothers often complain that their children are so well behaved only when they are not around" (3, p. 307).

The findings of the present study, as well as those of other related studies (1, 3, 6, 12), also have direct methodological implications. They emphasize the significance of the behavior of any adult *E* in a laboratory situation with children, pointing to the *social* nature of the play situations typically used in research with young children, and demonstrating that children's play responses are extremely sensitive to variations in the social setting of the play.

SUMMARY

The social play of pairs of young boys was observed in two sessions separated by two days, and the aggression of the older member of each pair was scored. Half the pairs' play sessions were in the presence of a permissive adult, and half were in the absence of any adult. Two-thirds of the Ss in the adult-present sessions were more aggressive in the second than in the first session, and all the Ss in the adult-absent sessions were less aggressive in the second than in the first session. This finding is in confirmation

ALBERTA ENGVALL SIEGEL and LYNETTE GAYLE KOHN

of the hypothesis, which was drawn from a consideration of the nature and effects of adult permissiveness with children and of the nature of young children's controls for aggression.

REFERENCES

1. HOLLENBERG, ELEANOR, & SPERRY, MARGARET. Some antecedents of aggression and effects of frustration in doll play. *Personality*, 1951, 1, 32-43.
2. LEVIN, H., & SEARS, R. R. Identification with parents as a determinant of doll play aggression. *Child Developm.*, 1956, 27, 135-153.
3. LEVIN, H., & TURGEON, VALERIE F. The influence of the mother's presence on children's doll play aggression. *J. abnorm. soc. Psychol.*, 1957, 55, 304-308.
4. MUNROE, RUTH L. *Schools of psychoanalytic thought*. New York: Dryden, 1955.
5. PHILLIPS, RUTH. Doll play as a function of the realism of the materials and the length of the experimental session. *Child Developm.*, 1945, 16, 123-143.
6. PINTLER, MARGARET H. Doll play as a function of experimenter-child interaction and initial organization of materials. *Child Developm.*, 1945, 16, 145-166.
7. REDL, F., & WINEMAN, D. *The aggressive child*. 1. *Children who hate*. Glencoe, Ill.: Free Press, 1957.
8. SEARS, PAULINE S. Doll play aggression in normal young children. *Psychol. Monogr.*, 1951, 65, No. 6.
9. SEARS, R. R. Influence of methodological factors on doll-play performance. *Child Developm.*, 1947, 18, 190-197.
10. SEARS, R. R., MACCOBY, ELEANOR E., & LEVIN, H. *Patterns of child rearing*. Evanston, Ill.: Row, Peterson, 1957.
11. SIEGEL, ALBERTA E. Film-mediated fantasy aggression and strength of aggressive drive. *Child Developm.*, 1956, 27, 365-378.
12. SIEGEL, ALBERTA E. Aggressive behavior of young children in the absence of an adult. *Child Developm.*, 1957, 28, 371-378.
13. SIEGEL, S. *Nonparametric statistics for the behavioral sciences*. New York: McGraw-Hill, 1956.

The first of these is the fact that the United States is a young nation, and that its history is a history of growth and development. It is a history of a people who have been able to overcome many difficulties and to build a great nation out of a small colony. The second fact is that the United States is a nation of immigrants, and that its history is a history of the struggle for the rights of these immigrants. The third fact is that the United States is a nation of free men, and that its history is a history of the struggle for the rights of these free men. The fourth fact is that the United States is a nation of law, and that its history is a history of the struggle for the rights of these laws. The fifth fact is that the United States is a nation of peace, and that its history is a history of the struggle for the rights of these peace-loving people. The sixth fact is that the United States is a nation of progress, and that its history is a history of the struggle for the rights of these progressive people. The seventh fact is that the United States is a nation of justice, and that its history is a history of the struggle for the rights of these just people. The eighth fact is that the United States is a nation of liberty, and that its history is a history of the struggle for the rights of these liberty-loving people. The ninth fact is that the United States is a nation of equality, and that its history is a history of the struggle for the rights of these equal people. The tenth fact is that the United States is a nation of unity, and that its history is a history of the struggle for the rights of these united people. The eleventh fact is that the United States is a nation of strength, and that its history is a history of the struggle for the rights of these strong people. The twelfth fact is that the United States is a nation of wisdom, and that its history is a history of the struggle for the rights of these wise people. The thirteenth fact is that the United States is a nation of courage, and that its history is a history of the struggle for the rights of these courageous people. The fourteenth fact is that the United States is a nation of faith, and that its history is a history of the struggle for the rights of these faithful people. The fifteenth fact is that the United States is a nation of hope, and that its history is a history of the struggle for the rights of these hopeful people. The sixteenth fact is that the United States is a nation of love, and that its history is a history of the struggle for the rights of these loving people. The seventeenth fact is that the United States is a nation of kindness, and that its history is a history of the struggle for the rights of these kind people. The eighteenth fact is that the United States is a nation of gentleness, and that its history is a history of the struggle for the rights of these gentle people. The nineteenth fact is that the United States is a nation of meekness, and that its history is a history of the struggle for the rights of these meek people. The twentieth fact is that the United States is a nation of mildness, and that its history is a history of the struggle for the rights of these mild people. The twenty-first fact is that the United States is a nation of lowliness, and that its history is a history of the struggle for the rights of these lowly people. The twenty-second fact is that the United States is a nation of modesty, and that its history is a history of the struggle for the rights of these modest people. The twenty-third fact is that the United States is a nation of humility, and that its history is a history of the struggle for the rights of these humble people. The twenty-fourth fact is that the United States is a nation of simplicity, and that its history is a history of the struggle for the rights of these simple people. The twenty-fifth fact is that the United States is a nation of plainness, and that its history is a history of the struggle for the rights of these plain people. The twenty-sixth fact is that the United States is a nation of plainness, and that its history is a history of the struggle for the rights of these plain people. The twenty-seventh fact is that the United States is a nation of plainness, and that its history is a history of the struggle for the rights of these plain people. The twenty-eighth fact is that the United States is a nation of plainness, and that its history is a history of the struggle for the rights of these plain people. The twenty-ninth fact is that the United States is a nation of plainness, and that its history is a history of the struggle for the rights of these plain people. The thirtieth fact is that the United States is a nation of plainness, and that its history is a history of the struggle for the rights of these plain people.

THE EFFECT OF VERBALIZATION IN CHILDREN'S LEARNING AS A FUNCTION OF CHRONOLOGICAL AGE

MORTON W. WEIR¹ and HAROLD W. STEVENSON*²

The University of Texas

Several recent studies have been concerned with the role of verbalization in children's learning. Children, as do adults (2), learn to discriminate among stimuli more rapidly if they are given pretraining experience in naming the stimuli (3, 4, 6, 9), or if they attach a verbal response to the stimuli during training (10, 11). Recall of stimuli has also been found to increase with prior naming of the stimuli (8). Significantly faster learning and a greater incidence of transposition have been found for verbal than for preverbal Ss (1, 7).

Although age differences in the effects of verbalization on learning have been noted (6, 9), possible changes in the effects of verbalization on discrimination learning with increasing CA have not been studied systematically. The purpose of this study is to investigate such effects with children of CAs 3 through 9 years.

One group of Ss at each age level is instructed to verbalize the name of the stimulus prior to each response, while another group is not instructed to verbalize in this manner. Several predictions about the performance of Ss are made. First, groups instructed to verbalize stimulus names should learn the discrimination more rapidly than Ss not given these instructions. This is in line with the assumption (12) that verbalization of stimulus names provides response-produced stimuli which increase the differences among the stimuli and aid learning by increasing the number of stimuli to which a discriminative response can become attached.

Second, in both verbal and nonverbal conditions learning rate should increase with increasing CA. Such an increase would be in accord with the assumption that learning ability improves with increasing age.

* Department of Psychology, University of Texas, Austin.

¹ This study is adapted from a master's thesis submitted to the Department of Psychology of the University of Texas by this author (14).

² The study was supported by a grant from the National Science Foundation (Grant 3280) to this author.

CHILD DEVELOPMENT

Finally, the difference in rate of learning between the verbal and non-verbal groups is predicted to decrease as CA increases. It is assumed (a) that as age increases Ss have stronger tendencies to verbalize implicitly the names of common stimuli, and (b) that such implicit verbalization provides response-produced stimuli similar to those occurring when Ss verbalize explicitly. Instructions to name the stimuli prior to responding should not be so helpful to S if S already has available the stimuli resulting from implicit verbalization. Older Ss, who are assumed to have stronger tendencies toward implicit verbalization than younger Ss, should be aided to a lesser degree than younger Ss by instructions to verbalize. As a consequence, the facilitative effects of instructions to verbalize should decrease with increasing CA.

METHOD

Subjects

The Ss consisted of 128 preschool and elementary school children selected on the basis of CA. A group of 32 Ss was selected at each of the following CA levels: 3-0 to 3-11, 5-0 to 5-11, 7-0 to 7-11, and 9-0 to 9-11. All of the preschool Ss in attendance at the schools visited and of the appropriate ages were tested. The elementary school Ss were selected at random from the class roll.³

Apparatus

The front of the apparatus was 17×37 in. and was painted gray. Protruding $\frac{1}{2}$ in. from the face of the apparatus were $10 \frac{1}{2} \times 3$ in. panels, with fronts of clear plastic bordered by $\frac{3}{4}$ -in. red plastic tape. Behind each panel was a 2×2 in. slide depicting an animal. The slides were visible to S only when illuminated from behind by 110 v., 6 watt bulbs. The slides were photographic negatives; thus, the animals appeared as white line drawings. The apparatus was constructed so that any of the five vertical pairs of slides could be illuminated on a particular trial. The S's task on each trial was to choose the correct member of the illuminated pair. When S pressed the correct panel, the lights behind both members of the pair automatically turned off. If S pressed the incorrect panel, the lights did not go off.

A pretraining apparatus was constructed in the same manner as the apparatus described above, except that it contained only one vertical pair of panels. The front of the apparatus was 17×10 in.

Common animals, such as a monkey, cat, and elephant, were depicted in the training slides. The animals depicted in the pretraining slides were

³ The writers wish to express their appreciation to the principals and the directors of the St. Louis School, Tarrytown Methodist Nursery School, All Saints Nursery School, Good Shepherd Nursery School, Austin High School Child Development Laboratory, and the University of Texas Nursery School for their cooperation in making Ss available for this study.

a bee and a pig. The rewards given at the end of the experiment consisted of a wide variety of balloons and small plastic toys.

Procedure

The *Ss* were obtained individually from the classroom and brought to the experimental room by *E*. The room contained a low table upon which the apparatus was placed. To the right of the apparatus was a second table with the pretraining apparatus.

Pretraining procedure. The *S* was seated in a chair facing the pretraining apparatus. The training apparatus was out of *S*'s view. The *E* stated that a new kind of game would be played and the pretraining apparatus was pointed out. The *E* said, "See these two boxes here. Well, you can push them like this." The *E* demonstrated how the panels could be pushed and then turned on the lights illuminating the pictures in both panels. The *S* was told, "Now you push on this one." (The *E* indicated the incorrect panel.) "See, nothing happens. Now try this one." (The *E* indicated the correct panel.) "See, the lights went out. Now that is what you are supposed to do. Every time two pictures come on you have to push on the one that you think will turn the pictures off. You always try to turn the pictures off on the very first push. Now let's try it again and see if you can turn the pictures off on the very first push." Pretraining was continued until *S* had made the correct response on three successive trials. The *E* concluded pretraining by saying, "It is always the same picture that turns the lights off. That's fine. Now this was just practice. Let's go over here where you can win some prizes," and took *S* to the chair in front of the training apparatus.

During the pretraining period the upper panel was correct for half the *Ss* and the lower panel was correct for the other half. The panel which was correct was randomized among *Ss*.

The *Ss* at each CA level were randomly divided into two groups of equal size prior to the pretraining period. One of these groups was designated the nonverbal group and the other the verbal group. During the pretraining period the procedure outlined above was followed with the nonverbal group. During pretraining *Ss* in the verbal group were asked to name the animals depicted when the panels were first illuminated. Since all *Ss* knew the names of the animals, it was unnecessary for *E* to identify them. A second difference in the procedure introduced for verbal *Ss* was that *E* requested *S* to name the animal depicted in the panel which they were going to press. The *E* told the verbal *Ss*, "Now remember, before you push on any picture you have to tell me the name of the picture."

Training procedure. After *S* had been seated in the chair facing the training apparatus, *E* said, "See this. Well, it is just like the game we practiced on only there are more pictures. I can turn on these two, or maybe these two, or any two I want." (The *E* pointed out the pairs of panels at random.) "And just like when we practiced, whenever two pictures come on you have to push the one which you think turns both pictures off." The

CHILD DEVELOPMENT

E informed *S* that if *S* learned to turn all the pictures off without making a mistake he could win any two prizes that he wanted. The *E* displayed the prizes that could be won. The *E* cautioned *S*s in the verbal group that they had to say the name of the picture before they pushed on it. As in pre-training, all *S*s were able to name the stimuli. The exact names were not required; for example, *S*s were allowed to substitute such words as "duck" for the goose. After every 25 trials *S* was told, "Remember, you have to turn the pictures off on the very first push."

A corrective procedure was employed. If *S* initially pressed the incorrect panel, he was allowed to make the correct choice. The criterion for learning was one run in which the correct panel in each of the five pairs was chosen first. If *S* did not reach this criterion in 100 trials, training was terminated.

The order of presentation of stimulus pairs was randomized in blocks of five trials so that each pair was presented once in each block. No pair of stimuli was ever presented twice in succession. A prearranged schedule of 28 successive blocks of five trials was constructed and the training trials for each *S* began with a block selected randomly from this schedule. In order to control for possible position or stimulus biases, two procedures were introduced. First, the panel in each pair which was to be correct was randomized separately for each *S* with the restriction that for an *S* no more than three upper or three lower panels were correct. Second, the slides were changed among the panels at random after every five *S*s had been tested. The *S*s at each age level were alternately assigned to the verbal and non-verbal conditions.

RESULTS

The average numbers of correct responses made in the verbal and non-verbal groups at each age level are presented in Figure 1. An analysis of variance reveals a significant difference in performance among the four age levels ($F = 2.96$, $df = 3, 120$, $p < .01$). The curves in Figure 1 do not, however, reveal a tendency for the average number of correct responses for either the verbal or nonverbal groups to increase consistently with increasing CA, as predicted. There is an increase in the average number of correct responses from CA 3 to 5 followed by a decrease from CA 5 through 9. An analysis taking into account the ordered character of the age variable (5) indicated a significant quadratic component in the over-all trend for the combined verbal and nonverbal groups ($F = 10.11$, $df = 1, 120$, $p < .01$). The linear and cubic components did not approach significance.

The verbal groups had on the average a greater number of correct responses than did the nonverbal groups. The analysis of variance indicates a significant difference between the verbal and nonverbal groups ($F = 14.37$, $df = 1, 120$, $p < .001$), which supports the hypothesis concerning the facilitative effects of verbalization on learning.

The prediction that the difference in average number of correct responses between the verbal and nonverbal groups would decrease with increasing CA was not supported by a significant interaction between verbalization and nonverbalization and CA ($F < 1$). A breakdown of the interaction term (5) does not reveal a significant linear, quadratic, or cubic component. As seen in Figure 1, such a tendency does emerge between CAs 3 and 5, but as CA increases to 7 and 9 the curves again diverge.

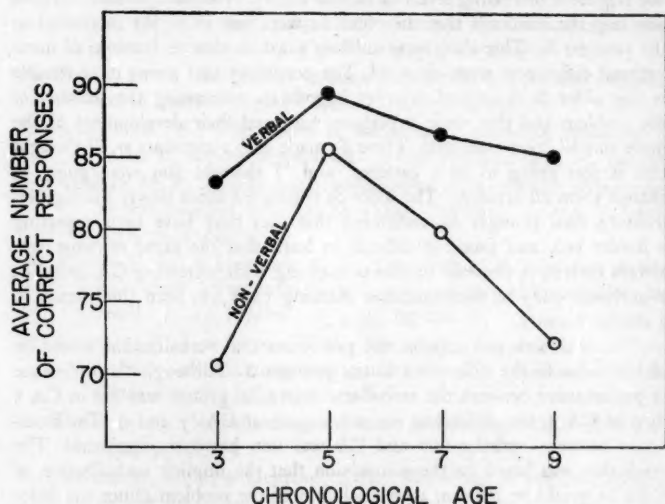


FIGURE 1—The average number of correct responses as a function of chronological age.

DISCUSSION

The prediction that instructions to verbalize would result in more rapid learning at all age levels was supported. This finding is in accord with the results of the studies cited earlier demonstrating that verbalization concerning the stimuli aids learning.

The prediction that rate of learning would increase with increasing CA was not supported. Although performance of the four age levels differed significantly, there was a decrease rather than an increase in rate of learning from CAs 5 to 9 in both verbal and nonverbal groups. Since these results are in contrast with a vast amount of evidence indicating that learning ability does improve with increasing CA, other variables must have been operating in this situation to provide the results that were obtained.

CHILD DEVELOPMENT

One possibility is that older Ss (CAs 7 and 9) came from a population different from that of the younger Ss. This does not seem likely, for the study was done at private schools attended by children of above-average ability. Children attending preschool are in general highly selected, and the results of group testing of the older Ss indicated they, too, were above average. The average IQ obtained for 28 of the 7-year Ss was 113.3 and for 28 of the 9-year Ss, 109.3. Thus, although the sampling of above-average children limits the generality of the results, the Ss at the various CA levels did represent increasing levels of mental ability. A second possible variable affecting the results is that the older Ss were not so highly motivated as the younger Ss. This also seems unlikely since no clear indications of motivational differences were observed. The possibility that seems most tenable is that older Ss developed complex hypotheses concerning the solution of the problem and that these hypotheses hindered their development of the more simple, correct solution. These Ss made such statements as, "I thought that it was going to be a pattern," and "I thought you were going to change them all around." The older Ss responded more slowly during pre-training than younger Ss, indicating that they may have been expecting a harder task and found it difficult to learn that the same response was always correct. A decrease in rate of learning with increasing CA found in a previous study of discrimination learning (13) has been interpreted in a similar manner.

The results do not support the prediction that verbalization would be of less value to the older than to the younger Ss. Although the difference in performance between the verbal and nonverbal groups was less at CA 5 than at CA 3, the difference increased again at CAs 7 and 9. The interaction between verbalization and CA was not, however, significant. The prediction was based on the assumption that the implicit verbalization of older Ss would be relevant to the solution of the problem. Since the older Ss in the nonverbal groups performed so poorly, it seems likely that, if implicit verbalization did occur, it hindered rather than aided the acquisition of the correct responses. The fact that explicit verbalization continued to aid Ss at the higher age levels in their performance indicates that verbalization of the names of the stimuli may have forced the older Ss to pay closer attention to the stimuli, thus increasing the relevance of the stimuli and decreasing the tendency of the older Ss to form complex hypotheses.

SUMMARY

Children at CAs 3, 5, 7, and 9 were trained to discriminate the correct member of five pairs of stimuli. Half of the Ss were instructed to name the stimulus prior to a response, while the other half learned the task without these instructions. Consistently better learning in the verbal groups and a curvilinear relationship between rate of learning and CA were obtained. The difference in rate of learning between the verbal and nonverbal groups was not found to vary significantly with CA.

REFERENCES

1. ALBERTS, E. & EHRENFREUD, D. Transposition in children as a function of age. *J. exp. Psychol.*, 1951, 41, 30-38.
2. ARNOULT, M. Stimulus predifferentiation: some generalizations and hypotheses. *Psychol. Bull.*, 1957, 54, 339-350.
3. CANTOR, G. N. Effects of three types of pretraining on discrimination learning in preschool children. *J. exp. Psychol.*, 1955, 49, 339-342.
4. DIETZE, D. The facilitating effect of words on discrimination and generalization. *J. exp. Psychol.*, 1955, 50, 255-260.
5. GRANT, D. A. Analysis of variance tests in the analysis and comparison of curves. *Psychol. Bull.*, 1956, 53, 141-154.
6. JEFFREY, W. E. The effects of verbal and non-verbal responses in mediating an instrumental act. *J. exp. Psychol.*, 1953, 45, 327-333.
7. KUENNE, M. R. Experimental investigation of the relation of language to transposition behavior in young children. *J. exp. Psychol.*, 1946, 36, 471-490.
8. KURTZ, K. H., & HOVLAND, C. I. The effect of verbalization during observation of stimulus objects upon accuracy of recognition and recall. *J. exp. Psychol.*, 1953, 45, 157-164.
9. NORCROSS, K. J., & SPIKER, C. C. The effects of type of stimulus pretraining on discrimination performance in preschool children. *Child Developm.*, 1957, 28, 79-84.
10. PYLES, M. K. Verbalization as a factor in learning. *Child Developm.*, 1932, 3, 108-113.
11. SCHAEFFER, M. S., & GERJUOY, I. R. The effect of stimulus naming on the discrimination learning of kindergarten children. *Child Developm.*, 1955, 26, 231-240.
12. SPIKER, C. C. Experiments with children on the hypothesis of acquired distinctiveness and equivalence of cues. *Child Developm.*, 1956, 27, 253-263.
13. STEVENSON, H. W., ISCOE, I., & MCCONNELL, C. A developmental study of transposition. *J. exp. Psychol.*, 1955, 49, 278-280.
14. WEIR, M. W. The effect of verbalization on children's learning as a function of chronological age. Unpublished master's thesis, Univ. of Texas, 1958.

POSITIVE REINFORCERS FOR EXPERIMENTAL STUDIES WITH CHILDREN—CONSUMABLES AND MANIPULATABLES

SIDNEY W. BIJOU* and PERSIS T. STURGES¹

University of Washington

This paper is an analysis of the kinds and the ways that positive reinforcers may be used in experimental studies with children. For the purpose of this presentation they have been classified as follows: (a) consumables, (b) manipulatables, (c) visual and auditory stimuli, (d) social stimuli, and (e) tokens. The first two categories will be considered here; the other three, in a paper to follow.

The material is from three sources: (a) papers published before 1947 dealing specifically with motivation in children; (b) experiments, technical notes, and theoretical analyses published between 1947 and the present; and (c) studies conducted at the University of Washington since 1954.² In regard to the second category, no attempt was made to summarize all experiments using positive reinforcers. Instead, studies were selected which illustrate typical practices.

The present discussion is limited to the relationship between positive reinforcers and instrumental or operant behavior. Let us spell out what we mean by the two key terms. By *positive reinforcers* we refer to those classes of stimuli which upon presentation strengthen the behavior that they follow. It cannot be said that stimuli referred to as "rewards," "incentives," or "goal objects" in studies with children are necessarily positive reinforcers. They may be for some subjects, not for others. The only way to find out is by submitting each to an experimental test: observe the behavior prior to the introduction of the stimulus and note the changes in behavior after presentation contingent upon a response. If the frequency of response is increased, then the stimulus is a reinforcer for that organism in that situation (18, 34, 36). By *operant behavior* we refer to those classes of responses that have effects upon the surrounding environment (34) and "for which eliciting stimuli have not necessarily been determined" (36, p. 22). We are therefore concerned with studies dealing with acquisition of sensorimotor skills, learning discriminations, formulating concepts, learning ab-

* Institute of Child Development, University of Washington, Seattle.

¹ Some of this material was presented at the 1957 meeting of the W. P. A. at Eugene, Oregon, by the junior author.

² Research reported here was supported by the National Science Foundation (G-2141).

CHILD DEVELOPMENT

stractions, solving problems, or responding to schedules of reinforcement. All or most of these problems have been treated in the laboratory with some variation of multiple-choice and free operant techniques.

In discussing consumables and manipulables, attention is given to some of the conditions influencing their effectiveness. These include the age of the child and such situational factors as the kind of experimental task, the instructions about the nature of the task and about disposition of the reinforcers, the method of presenting the reinforcers, and the behavior of the adult in the experimental room. Of course, historical conditions also play a part. To take them into account, however, would be to go beyond the limited objective of this analysis.

CONSUMABLES

Procedures in which consumables, edibles and drinkables, are contingent upon an instrumental act are discussed here. Those in which the consumables are delivered after completion of a series of responses linked with tangible conditioned reinforcers will be treated in a discussion of tokens in a paper to follow.

Edibles

Edibles, of various kinds, frequently have been used as positive reinforcers in both multiple-choice and free operant experimental studies with children. These procedures will be discussed in the following order: (a) experimental studies with preschool children, (b) those with older children, and (c) those with deviant children.

Preschool children. Multiple-choice experiments are particularly suited for the use of edible reinforcers and may either give the child freedom to do as he pleases with the edibles or instruct him to save them until the experiment is over. A study by Hunter on transposition behavior (16) is an example of a multiple-choice procedure with preschool children in which S^8 may either eat or save what is received. The question at issue was whether behavior on a transfer task is determined by the absolute sizes or the relative differences in size of pairs of circles. Three experiments were performed with two distinctively different procedures. In the first and second, with raisins as reinforcers and S s from 4-11 to 5-3 and 5-0 to 5-11 respectively, E described the task as a game and explained the rules:

Each time, you have to look at both cards and see which is right for, if you pull out the drawer—look, there's a drawer here and another there—beneath the right card, you get a raisin. You like raisins? If you get a raisin you can keep it or eat it or do what you like with it. If you pull out the drawer under the wrong card, there will be no raisin there and you won't get any. Now let's pull out the drawer under this card. (Pull out drawer

⁸ In describing procedures the subject will be designated as S and age of subject will be expressed by two figures indicating years and months. Thus, the age of a 3½-year-old child will be noted as 3-6. The experimenter will be referred to as E .

beneath wrong card.) There's nothing there so that must be the wrong card. Now let's pull out the drawer under this card. (Pull out drawer beneath right card.) Oh, look, there's the raisin. Take it (16, pp. 117-118).

The first time the "right" drawer was opened *S* took the raisin. The first time the "wrong" drawer was opened the experimenter, who was seated at the opposite end of the table, "held the other drawer from behind so that the subject could not open it, and informed the subject that he must open the 'right' drawer the first time, otherwise he could not get the raisin," (16, p. 117). From then on *E* arranged the cards for the trials, "baited" the drawers, kept records, and held fast to the "right" drawer on errors to prevent spontaneous corrections.

In the third experiment, materials and instructions were tailored for *Ss* between ages 1-2 and 2-8. *E* sat on a low stool on the opposite side of a table from *S* with a stack of blocks and a sack of currants on the floor by his side to be used as reinforcers. At the beginning of each trial a currant was placed under a block of appropriate size and pushed within reach of *S*. Correction was sometimes allowed and sometimes not, at the experimenter's discretion, "since a rigid scheme was not found optimal for learning" (p. 124). Furthermore, since some of the blocks were difficult for the *Ss* to manipulate, *E* now and then gave a little unobtrusive aid.

These procedures include a considerable amount of "flexible" adult-child interaction. Perhaps in working with 1- and 2-year-olds, a certain amount of guidance is necessary to help the child "in the situation"; however, it is difficult to evaluate the relative influence of the edibles and the social interactions, verbal and otherwise. It would facilitate analysis if variations followed a definite plan that could be unambiguously described (as for example, successive approximations with objective behavioral criteria).

Multiple-choice experiments may be arranged so that *S* sees and even handles the edible reinforcers, but is instructed to save them until the end of the session. Let us describe two studies with this variation. In the first, by Alberts and Ehrenfreund (1), *Ss* were required to discriminate white squares on a black background. One pair of squares at a time was presented to *Ss* from 3-0 to 3-11 and 4-5 to 5-6, from the State College of Washington Laboratory Nursery School. The response was opening compartment doors. Correct choices resulted in finding a gumdrop. *S* was brought into the room after a "warming-up" procedure with a toy dog, and was given instructions by *E* who remained behind the apparatus panel where he arranged the problems and recorded responses. "When a trial was successful (i.e., *S* opened the smaller door), *S* took the gumdrop and placed it upon the napkin which had been spread on another small chair nearby, while *E* closed the door and the curtains" (1, p. 33). Since some *Ss* reacted emotionally to nonreinforcement, positive social stimulation followed an incorrect response. "*S* was assured of another opportunity to find the candy. . . ."

In the second study, Brackbill and O'Hara (8) were interested in comparing the relative effectiveness of reward and punishment for discrimina-

CHILD DEVELOPMENT

tion learning. An M&M candy was given for a correct response and one was taken away for an error. The stimuli were three boxes differing in color and size. The task, for 43 boys with a mean age of 5-10, was to learn a simple position alternation sequence between the first and second boxes. At the beginning of the session *S*, who sat opposite *E* at a small table, was "staked" to a box of M&Ms, after which *E* said to the *Ss* in the reward-punishment group:

Now, I have something very special for you to do, and if you can do it, you can have some more of the candy. I'm going to put one candy at a time under one of these boxes. And every time you pick up the box that has the candy under it, you can keep the candy. But, every time you pick up the WRONG box, you'll have to GIVE BACK one of your candies. (*E* pointed to *S*'s pile.) Understand? Now! THERE IS A WAY TO FIND THE RIGHT BOX EVERY SINGLE TIME. SEE IF YOU CAN FIND THE WAY (8, p. 3).

The same instructions were read to the *Ss* of the reward-only group, except that nothing was mentioned about withdrawing an M&M after a wrong response. Following the first correct response, *E* said to all *Ss*: "Now every time you get a candy, pick it up and put it down there, and all the candy that you put there will be for you to keep." Following the first incorrect response for the reward-punishment group, *E* said: "That was the wrong box, so you'll have to give back one of your candies." At the end of the session *E* put the M&Ms in *S*'s box into a sack and gave it to him.

In this, as well as in the Alberts and Ehrenfreund study, *E* played an active role in dealing with the reinforcers and *S*'s reactions to various parts of the experiment. Because of such interaction it is difficult to determine the effectiveness of the candy relative to the social reinforcement that may have been involved. There is no objection to the use of social reinforcement. The only question is which set or sets are operating in the experiment. The challenge for future research is to carry out the procedures required under circumstances that are explicit regarding the reinforcements in operation.

The free operant technique is, of course, a natural for edible reinforcers, and in recent years an increasing number of such studies with children have appeared. A terse statement of the salient characteristics of the method is given by Ferster: "The use of the free operant is a method of wide generality; it refers to any apparatus that generates a response which takes a short time to occur and leaves the animal in the same place ready to respond again. The free operant is used in experiments when the main dependent variable is the frequency of occurrence of behavior" (12, p. 263). Three examples with preschool children will be considered here. Each has a bearing on some of the factors influencing the use of edibles as reinforcers. In the first, Warren and Brown (37) used hemispherical pellets of candy made of sugar, corn syrup, and gelatin ("Little Gems"), delivered by a vending-type machine modified for lever pressing. *Ss*, 2-0 to 5-0 in age, performed alone in the experimental room, decorated and arranged for

preschool children. The preliminary part of the procedure was described as follows:

For each experimental period one of us brings the child individually from his house to the psychological laboratory, enters the playroom with him, and if necessary plays with him for a few minutes. Then the subject is locked in the room with a toy, book, and lever. The toy and book generally occupy the child's interest when he is not at the lever. . . . If for any reason the subject becomes overexcited, the experimenter enters the playroom. If necessary, the experiment is stopped and continued the next day (37, p. 183).

Each *S* was seen once a day for 22 minutes, and each, serving as his own control, was observed under five conditions (operant conditions, extinction, spontaneous recovery, disinhibition and periodic reconditioning) over 15 experimental days. Strength of drive for "Little Gems" was measured by behavior. *Ss* who ate all the candy while in the playroom or shortly thereafter were rated as having strong drive; those who ate little or no candy or objected to receiving candy were judged as having weak drive, and those who behaved somewhat in between were classified as having medium drive. One of the findings was that "the strength of the child's appetite for candy reward used in this experiment determines the orderliness of conditioning, extinction, and periodic reconditioning" (37, p. 203).

The second investigation, by Siegel and Foshee (33), was concerned with the relationship between number of reinforcements and rate of extinction. *Ss* from ages 2-11 to 3-8 were also required to press a bar for cylindrical-shaped party mints, each weighing approximately .2 of a gram. One person stayed with the child throughout, "making for a uniform social environment." He attempted to keep his behavior neutral and constant. Observations, made two to three hours after breakfast, were initiated by inviting *S* to come and see a new game.

If he indicated acceptance, he was led to a nearby experimental room and told that he could sit or stand before the apparatus. The *E* then pointed to a single pellet in the tray and asked, "What is it?" If *S* indicated recognition, he was told that he could eat it. If no recognition was apparent, *E* told the child that it was candy and that he was free to eat it. Next, *E* pointed to the lever and inquired, "What do you think would happen if you pushed this?" If *S* pressed the lever, ate the candy, and pressed again, *E* said nothing further. If *S* pressed the lever and then failed to eat the candy, *E* said, "You may have as much as you like." No further comment was offered (33, p. 13).

The third study, by Baer (3), resembled Warren and Brown's in that it aimed to explore operant processes, and Siegel and Foshee's in that an adult was in the experimental room. *Ss*, from 4-0 to 4-6, were attending the University of Chicago Nursery School. The task was to press down on a wooden door knob for M&M candies delivered in paper cups by a revolving dispenser modeled after a lazy-susan. After orienting *S* to the situation and to the task as a game, *E* waited until *S* responded to the first candy,

CHILD DEVELOPMENT

and then suggested that he could eat it if he wished. (All *Ss* ate them immediately.) The notation on the relationship between *E*'s behavior and the *S*'s performance during the period of nonreinforcement is of particular interest. "The extinction rate apparently was affected by a grossly defined social condition in which the experimenter injected his attention and approval into the extinction process" (3, p. 17).

In these three studies, it is apparent that in addition to candy *per se*, other factors were influential. First, there were differences in kinds of candy. It will be recalled that Warren and Brown found that differences in strength of appetite for candy influenced the orderliness of performance. Different kinds of candy assuredly have differential appeal and should also be expected to produce variations in performance. Second, the activities of the adult in the experiment unquestionably influenced performance. Baer's finding of the effect of attention and approval on extinction adds to the accumulating evidence. His findings are particularly significant since he did not make his attention and approval contingent upon a response although *S* may have so interpreted it. Third, differences in instructions—statements to induce *S* to participate and orient him to the experiment task—must have made for differences in performance. Although no studies have been cited here to demonstrate this, a large number of investigations on adult *Ss* (especially those using "projective techniques") indicate that it is a safe assumption (32).

Older children. Let us now consider two studies with older children showing how instructions and social participation may be drastically reduced. The first is by Lasko (23), illustrating a simple but ingenious method of studying intermittent reinforcement. *Ss*, 11-year-old school children, were presented with an apparatus consisting of a box with 96 holes across the front panel. They were to start at one end and insert a stick into the holes. The stick ejected an individually-wrapped, penny Hershey bar. One group was reinforced after each response; the other group, after every other response. Little or no social action was necessary.

The second investigation, by Azrin and Lindsley, was concerned with cooperative behavior (2). *Ss*, ranging in age from 7 to 12 and drawn from two settlement houses in Boston, were brought to the experimental room, two at a time. They sat at long ends of a table 6 ft. in length and 3 ft. in width. Physical contact between the children was prevented by a wire screen mounted on the table. Each *S* had a stylus which could be inserted in any of three holes in the table. *E* gave the following instructions and then left the room.

This is a game. You can play the game any way you want to or do anything else that you want to do. This is how the game works: Put both sticks (styli) into all three of the holes. (This sentence was repeated until both styli had been placed in the three available holes.) While you are in the room some of these (the experimenter, *E*, held out several jelly beans) will drop into this cup. You can eat them if you want to or you can take them home with you (2, p. 100).

The Ss had to discover that, if one of them inserted his stylus in one of the holes on his side of the table within 0.04 sec. after the other had inserted his stylus in a comparable hole on his side, a red light flashed and a jelly bean tumbled down a chute into a tray accessible to both Ss. Azrin and Lindsley stated that instructions were kept brief so that the initial acquisition of cooperative behavior could be studied; the procedures could be used with subjects having difficulty in understanding instructions, such as the retarded or disturbed; and the effect of instructions on behavior would be minimized.

Deviant children. Let us now turn attention to the use of edibles with deviant children. As has been mentioned earlier, in a multiple-choice type of instrumentation there is a tendency to combine edibles with social factors. Here are two examples with deviant children. A study by House, Zeaman, Orlando, and Fischer (15) was concerned with discrimination learning and transfer in institutionalized imbeciles. Stimuli varying in form were presented two at a time by means of the Wisconsin General Apparatus Test. Correct responses were rewarded by an M&M candy in the hole beneath the stimulus object and the word "good" delivered by E. Incorrect responses were followed by a "No." For the most part, the correctional method was used and S could do as he pleased with the candy. E operated from behind a one-way screen. He arranged the stimuli for each presentation, pushed the tray to S's side of the screen for a trial, and pulled it back after a response. The procedure relevant to reinforcers is as follows:

On the subject's initial exposure to the experimental situation, two or three pretraining trials were given as demonstration with a minimum of verbal explanation. The subject was first asked if he liked candy and was given a small piece. If he responded positively to the candy (i.e., either ate or hoarded it), he was told that he was going to play a game for candy. For the first presentation of the stimulus tray, both food cups were left uncovered with candy placed in one of them. The experimenter asked, "Can you find the candy?" and pointed it out if the subject failed to see it. On the next trial, a grey wedge was placed over the food cup containing the candy, with the other cup left uncovered. Again S was asked to find the candy and aided by the experimenter if he failed to move the wedge. Another trial with the wedge was given at this point if the subject seemed slow. At this point, discrimination trials were begun (15, pp. 6-7).

In another multiple-choice study on the institutionalized retarded, Cantor and Hotell (11) were interested in the effect of amount of reinforcement on rate of learning. E explained and demonstrated a form discrimination task and arranged the problems by placing one peanut or four peanuts under the correct stimulus box. "When S made a correct choice, he was allowed to eat the food reward; when an incorrect choice was made, S was not allowed to lift the correct box, thus being deprived of the reward." Each of the authors served as E for half the Ss. The performance of the low and higher reward groups was not significantly different, but there was a significant interaction (at the 5 per cent level) of rewards by experimenters,

CHILD DEVELOPMENT

indicating that the behavior of *E* had some control over the *S*'s performance.

The next two examples on deviant children are of the free operant type. They suggest, at least with this type of method, that it is not essential to compound edibles with social reinforcers. The first one is part of a project on psychotic behavior by Lindsley (25). The experimental situation was a specially constructed cubicle in which *S*, working alone, was required to pull a plunger to obtain reinforcers delivered automatically. The reinforcers consisted of a mixture of candies including jelly-beans, corn candies, gumdrops, sour-balls, peanuts, chiclets, M&Ms, small Hershey bars, and Tootsie Rolls. Since the *S*s, ranging in age from 7 to 14, would not stay in the room alone at first, they were oriented and given preliminary training in gradual stages similar to Warren and Brown's free operant technique. Of special interest here is the attempt to evaluate the effect of food deprivation on performance. The report states:

In our usual experimental procedure the patients come to the experimental rooms an average of two hours after they have eaten. In order to obtain an exploratory determination of the effects of food deprivation on candy-reinforced behavior, three patients missed breakfast before the experimental session. These children showed no significant changes in the number of responses per hour as a result of this 20-hour food deprivation. These preliminary results suggest that short periods of food deprivation will not significantly increase the rate of responding for candy reinforcement (25, p. 5).

A second research program by House, Zeaman, Orlando, and Fischer (15) was on the developmentally retarded and was patterned after Lindsley's. Major interest was on the process of operant discrimination; hence, the apparatus was modified so that *S* faced a panel with two windows for presenting stimuli and a plunger for responding under each. Presentation of stimuli and delivery of reinforcers were controlled by electric devices. "The reinforcers are small, commercially available candies including: miniature marshmallows, candy corn, M&Ms, and chocolate chips (the kind that go into Toll-House cookies)" (15, p. 87).

Drinkables

There is no reason why drinkables should not serve as effective reinforcers. Small amounts of juice or milk can be delivered in sealed sanitary cups, or squirted into a stationary cup or glass. Yet, only two investigations using some kind of drink have come to our attention. For example, under Notes and Discussion in the 1949 volume of the *American Journal of Psychology*, Fuller reports a provocative investigation of operant conditioning of arm movement in an 18-year-old "vegetative idiot" using a warm milk-sugar solution (13). Observations were conducted before breakfast so that *S* was without food for 15 hours prior to training. Although the attending physicians at the institution thought that the subject could not learn anything, the response of arm movement was conditioned in four experimental sessions. The milk was delivered directly to *S*'s mouth. In

addition to demonstrating the feasibility of drinkable reinforcers, the study suggests that we have yet to apply known learning principles to the training of the retarded in a thorough and systematic way.

Summary

Consumables, edibles and drinkables, have been discussed as reinforcers in experiments with children. Edibles were considered first. A large variety of edibles has been used effectively with children in both multiple-choice and free operant tasks. Procedures involving multiple-choice tasks may either give the child freedom to do as he pleases with the edibles, or instruct him to save them until the experiment is over. Three experiments on preschool children were cited in illustrating the first practice, and two the second. It was pointed out that all involved the experimenter in the procedures, thereby making it difficult to evaluate the reinforcing function of the edibles as such. Procedures following free operant techniques almost have to allow the child to do as he pleases with the reinforcers. Procedures from three studies on preschool children were described, and it was noted that kind of candy, instructions to the child, and the behavior of the experimenter influenced the behavior of the child. Six additional studies were cited: one restricted operant and one free operant showing techniques with middle childhood children, two multiple-choice types on retarded children, one free operant on retarded subjects, and one free operant on emotionally disturbed youngsters.

Mixtures of edibles are most serviceable in studies requiring the child to come to the laboratory for many sessions. Mixtures of edibles and manipulables will be discussed in the following section. Drinkables were treated only briefly. Although drinkables may serve as reinforcers only a few studies have used them. One was cited involving a grossly retarded subject.

Findings from this analysis and from experiences at this Laboratory have been brought together in the form of practical suggestions for the use of consumables as reinforcers:

1. Prior to initiating a study with consumables, it is suggested that parents, teachers, principals, and/or superintendents be briefed on the kind and amount of consumables each child will receive from participation in the investigation. In most situations the information will be gratefully received and cooperation will be enhanced. In some cases, parents may express reservation on the basis that sweets are detrimental to the child's teeth or health and interfere with his eating routines, and teachers may add that foods brought from laboratory tend to "mess up" the class room and school equipment. Under such circumstances it is obviously advisable to plan on alternate reinforcers such as toys, trinkets, hobby items, music, movies, or social interaction.

2. Select the type of consumable that is acceptable to parents, teachers, or other adults responsible for the children, as well as advantageous to the purpose of the study and to the ease of reporting and replicating. It is suggested that the reinforcers have low calories, be individually wrapped, or

CHILD DEVELOPMENT

have a hard consistency or at least a hard surface, so that they will not readily stick, melt, or crumble.⁴ It is also recommended that consumables be small (about half the size of jelly beans) and uniform in size, and that they be easily identifiable and readily reproducible or obtainable. Mixtures of candies with tiny cookies, Trix cereal, nuts, and raisins are advocated for studies requiring long periods of observations or a series of observations extending over several months.

3. To reduce the influence of uncontrolled reinforcing stimuli, it is suggested that minimum instructions be used, that all instructions be given in a standardized manner, that reinforcers be presented by mechanical or electronic means—unless social reinforcers are also experimental variables—and that, if an adult must be present in the experimental situation, his behavior be evaluated experimentally or statistically.

4. To encourage investigations aimed at duplicating, elaborating, and extended findings, it is suggested that a complete and detailed account be given of the consumables used, the instructions employed, the method of dispensation, and the controls exercised over other social and physical stimuli.

MANIPULATABLES

Manipulables refer to toys, trinkets, and hobby items—objects which apparently derive their reinforcing properties from tactile, visual, and auditory stimulation. Some of the practices with manipulables will be considered here, others in the section on tokens in the paper to follow.

Toys

Toys—the five-and-ten-cent-store variety as well as the more expensive kinds—may serve as reinforcers in at least three ways: (a) presented for play for short periods during the experiment, (b) presented for play during the experiment and given to keep at the end, and (c) displayed during the experiment and given to keep at the end.

Limited periods of toy play as reinforcement are exemplified in a study on inferential behavior by Kendler and Kendler, and on concept learning by Roberts. In both, preschoolers were the Ss. The Kendler and Kendler experiment (19) used Ss from a private nursery school, ranging in age from 2-10 to 4-11. In the preliminary phase of the study *S* was required to pull a ribbon to obtain a small stuffed red ladybug, and a gold chain for a small stuffed gray chicken. "Although a string was clipped to both sub-goals to prevent them from being pulled more than a few inches past the apparatus, there was sufficient play in it to allow *S* to handle the stuffed toy. After a few moments of handling, the ladybug was retracted by *E*" (19, p. 312). *S* was then asked to pull the ribbon again. This time it not only brought forth the ladybug but also a toy foreign sports car. The child

⁴ The Kirkman Pharmacal Co., Seattle, Wash., is exploring possibilities of a standard experimental sugarless candy in attractive colors with fruit flavors.

"was permitted to play with the car about one minute. During this time he made it go around the room once and tooted the horn two or three times" (19, p. 312). *E* took an active role throughout.

In the Roberts' study (31) the *Ss* were from 2-0 to 5-9 in age. Each was shown a two-story gray wooden house with a roof made of glass. The house had three different colored doors, and the upper part of the door had a ridge upon which an airplane (the same color as the correct door) was sitting and which could be triggered so that the airplane would fall off. *E*, who remained with the child, gave the following instructions: "This house has three doors; all of them open, but only one door makes the airplane fall. You open one door and see if you can make an airplane fall" (31, p. 120). If *S* opened the correct door, he could play with the airplane while *E* completed his records of the response and prepared for the next trial. If he opened the wrong door, *E* said: "No, that door doesn't make an airplane fall. See, this is the one. This door makes an airplane fall." She then opened it, but the child could not play with the airplane since it and the house were removed at once. However, he could play with the other toys in the room until the next trial. As in the Kendler and Kendler study, *E* took an active role in the proceedings.

Giving youngsters toys to play with during the session and to keep afterwards was the practice followed by Stevenson (35) in a study on latent learning and by Lewis (24) on a problem dealing with "intermittent reinforcement." In the first study *Ss*, who were from a university housing project, ranged in age from 3-0 to 6-0. After preliminary training in opening locks, each was required to run down one side of a large V-shaped "maze" and open a box with a key. On the reinforced trials, *S* found a bird-, a flower-, or an animal-sticker. "... *E* asked him to bring the reward to the starting point and paste it on a large placard placed there. As the child pasted, the assistant replaced the lock, key and reward" (35, p. 18). As is apparent, *E* and his assistant were actively involved throughout.

Stevenson noted that the behavior of the older *Ss* (ages 5 and 6) appeared to be directed toward obtaining the reinforcements and complying with *E*'s requests, and their performance showed a consistency and directedness that was missing in many of the younger *Ss*. The younger *Ss* (ages 3 and 4) seemed to regard the reinforcement as but one of a number of interesting aspects of "the game" and showed considerably more manipulation of the irrelevant objects. A similar observation on differences in behavior of younger and older preschoolers in relation to trinkets will be discussed in the next section.

In the Lewis study (24) each *S* was given 20 plastic toys (cowboys, football players, and the like) at the beginning of the session and was told that it was possible to win many more but that it was possible to lose all of them, too. The youngest child in the group was 6-5, the oldest 7-5. Each subject was presented with four electrical switches and told that, if the correct one were pushed, a red light would flash and he would receive an additional toy; if the wrong one was pressed, a blue light would flash and

CHILD DEVELOPMENT

a toy would be taken away. The lights were controlled by *E* in order to distribute the reinforcers in predetermined percentages. At the end of the session the subject was told that he could keep the toys remaining. As in the other three studies, *E* took an active part in giving instructions and in presenting and withdrawing the toys. Incidentally, it may be recalled that Brackbill and O'Hara used a similar procedure with M&M candy to study the relative effectiveness of reinforcement and punishment in discrimination learning (8). Was the Brackbill and O'Hara study concerned with intermittent reinforcement or was the Lewis study concerned with reinforcement and punishment? According to the definition of terms ordinarily used in an objective behavioral analysis (34, 36), Lewis must have been dealing with reinforcement and punishment—in a complicated situation.

In a study of transposition behavior with *Ss* 2-6 to 5-10, Kuenne (22) used a variation of the above procedure. The child was given toys to handle and play with during the session and was given his choice of one of the toys to keep. Each *S* was seated in front of a panel with two hinged lids, behind each of which was a box. "In the event of a correct response, the box opened and a toy was found inside, while in the event of an incorrect response, the box was found to be locked" (22, p. 478). Colorful wood and cardboard toys of the five-and-dime variety were used and were collected in a cardboard box which the child held on his lap. At the conclusion of the experiment the box of toys was turned back to the *E* and the child was offered one selection from the cardboard toys to take home.

The third procedure with toys—presented contingent upon a response yet unavailable for possession or play until the end of a session—requires some kind of a dispensing apparatus and a clear plastic box to serve as a "showcase." With such an arrangement toys may be presented according to experimental plan, be seen during the experiment, and be given to keep at the end of the session. Studies following this procedure will be described in the section on trinkets.

Trinkets

One of the earliest notes on the effectiveness of trinkets as reinforcers is in a review of child development for 1952 by Nowlis and Nowlis (26). Since then multicolor and multiform plastic and metal trinkets, or charms, as they are sometimes called, have served in studies with boys and girls from 2 to 12. In some, they have been delivered but were not available for handling until the end of the experiment; in others, they were immediately available.

An example of the "you-can-see-but-can't-touch" method with nursery school children was described by Bijou (4). The apparatus was a wooden box with two holes, one above the other. When a ball was dropped into the upper hole, it rolled through a series of tubes and came out the lower hole. (Since the child must wait 3.3 sec. for the ball to return to make the next response, the situation is a restricted rather than a free operant.) To the left of the bottom opening was a clear plastic box with a closed cover

to receive trinkets. In preparing for the study, the *E*, a young woman, first made herself known to the children by watching them at play in the yard and in the school rooms. This phase ended when the children began to regard her as a member of the nursery school staff. The study got under way when the first child was selected to come and "play with some toys." The preliminary experimental procedure was as follows:

On entering the laboratory, she suggests that the child sit at a table upon which are a plastic dog and a tin dog house. After a few introductory remarks, the experimenter puts the dog in the house and shows the child that tapping on the table will make the dog shoot out. She next demonstrates that the dog will also pop out if called by his name, "Sparky." The subject is then encouraged "to make the dog come out." Once the child takes charge of the toy, play is permitted until he has made the dog come out at least four times. A delay of 15 sec. or more between insertion and pop-out is the criterion for termination of warm-up activity (4, pp. 165-166).

The child was then invited to sit in front of the table with the apparatus. A plywood screen was removed and *E* followed this procedure:

She shows a handful of trinkets and says: "You can get some of these to keep, to take home with you." She hands the child the ball and says: "When you put the ball in there (pointing to the upper opening) some of these toys will come down here (pointing to the covered plastic box). The ball will come out here again. You can put it back into the top hole as many times as you like. We will leave all the toys that come down in the box until we are finished, and are ready to go." The experimenter takes a seat six feet to the rear of the subject and offers no further suggestions or instructions. Comments and questions addressed to her are reacted to in a pleasant non-directive fashion, e.g., "You think the thing is broken." "You think no more will come out." When the experiment has been completed the receiving box is opened, the trinkets are given to the child, and the child is returned to his group (4, p. 166).

This procedure was used in two studies on intermittent reinforcement (5, 29) and one on the relationship between strength of response and number of reinforcements (30). A variation was employed by Kogan (21) in an analysis of presenting and withdrawing trinkets. *Ss* ranging in age from 3-0 to 5-11 were drawn from a university housing unit. To reinforce a response, a trinket appeared in a plexiglas window suspended from a hook on a horizontal chain which moved the trinket into view. To remove a reinforcer, chain movement was reversed causing a trinket to disappear from sight. At the end of the session *S* was given the trinkets displayed in the window. It was hoped that this procedure would eliminate the need for an adult to present and withdraw reinforcers as in the Brackbill and O'Hara and the Lewis studies. The device was not entirely satisfactory, however, since considerable training was required before *S* learned that every time one moved out of sight it was his loss. Further investigations have produced other devices and procedures which accomplish this purpose in a more efficient manner.

CHILD DEVELOPMENT

An example of a free operant method in which trinkets were available following a response has also been described by Bijou (6). The apparatus was a box with a lever for responding and a tray for receiving trinkets. The initial part of the procedure followed the practice of having an adult become acquainted with the children in the nursery school and then inviting them in, one at a time, to the laboratory to play games and get some trinkets. The latter part of the procedure was as follows:

On entering the laboratory, she closes the door, takes him to the middle of the room, and says, "You can play with the toys as long as you like." She disappears behind the cardboard screen, sits down, clips the ear-piece on her left ear (away from the side where the child would see it if he were to approach), and pretends to be reading or writing. During the experiment, if the youngster leaves the "toys" to peek around the screen, she looks at him pleasantly. If the child comes to talk about the trinkets, toys or anything else, she also reacts in a pleasant manner but says nothing. In response to direct questions she says, "You can play with the toys as long as you like." This is all she is permitted to say. All her other behavior toward the child is prescribed by *E* and communicated to her through the ear-piece (6, p. 247).

In one of the studies with this procedure interest was focused on determining the minimum number of trinkets necessary to obtain stable performance within a session and between sessions. The *Ss*, from the University of Washington Nursery School, ranged in age from 2-6 to 4-11. In each session seven trinkets were dispensed but 19 reinforcements in all were given. For some *Ss*, the remaining reinforcers were pieces of brightly colored paper; for others, they were simply the humming sound of the dispenser motor. Interval and ratio schedules were set up so that the first two reinforcements were trinkets and the interval between trinkets increased progressively. Thus trinkets were dispensed on the 1st, 2nd, 4th, 7th, 11th, 15th, and 19th reinforcement. As can be seen, the blocks of responses between trinkets were 0, 1, 2, 3, 3, 3. With both trinkets and hum, and trinkets and paper, a relatively steady rate was maintained by only one from the junior group (ages 2-6 to 3-3) and by five *Ss* from the senior group (3-3 to 4-6), out of a total of 25. For the remaining 19 children, the sessions became shorter, response rates decreased, and there were frequent refusals to return. Actually this was a double schedule—the distribution of reinforcements on an interval or ratio schedule, and the distribution of trinkets among the reinforcements on an increasing ratio basis (7). As will be discussed in the next section, mixtures of edibles and trinkets have resulted in stable performance over a number of sessions.

Casual observations in this and in other studies using the free operant method indicate age differences in response to trinkets, comparable to those with stickers described by Stevenson (35). Performance of the *Ss* in the older groups (ages 3-3 to 4-6) gave strong indication that they were "working for the trinkets," since they spent relatively little time with the other toys in the room. For the *Ss* in the younger group (age 2-6 to 3-3) appar-

ently the trinket was only another interesting aspect of the total situation, and they often spent as much time playing with the other toy in the situation as they did with the manipulandum.

Trinkets may, of course, be made immediately available in multiple-choice situations. Such a procedure was used by Calvin and his associates in a series of discrimination studies (9, 10, 27) on normal children and by Plenderleith on retarded children (28). In one of the Calvin studies (10) interest was centered on the effect of nondifferentiated reward and non-reinforcement on discriminative learning in 10 Ss 10-0 to 11-0. *S* was presented with two identical plain opaque red plastic cups, mouth down and 4 in. apart. Two $3 \times 4\frac{1}{2}$ in. white cardboard cards which served as cues were placed in horizontal positions, one in front of each cup. The procedure was described as follows:

If *S* pointed to the cup with the reward (a small plastic toy of the kind given in gum-ball machines) under it, *E* picked it up and gave the reward to *S*; whereas if *S* pointed to the cup with nothing in it, *E* first picked up the chosen cup to show that it had not covered a reward and then immediately lowered the screen to terminate the trial without giving *S* a reward (10, pp. 439-440).

The Plenderleith study (28) was a test of Lewin's concept of rigidity as a personality trait of the retarded. Normal and retarded Ss ranging in mental age from 5-6 to 5-11 were observed in discrimination learning tasks involving pairs of pictures. The procedure of orienting the *S* to the situation, giving instructions, setting up the choices, dealing with behavior arising from nonreinforcement and other conditions, was as follows:

The stimuli used in this study were a series of pairs of pictures. No two pictures were the same. The pictures were inserted into slots in a movable frame behind which was a concealed reward trough. The child was instructed to select one of the two pictures by pushing the card, and if the correct card was chosen the child found a reward (small charms purchased at novelty counters) in the trough behind it. If the wrong card was chosen, no reward was found. The noncorrection method was used throughout (28, p. 108).

In both multiple-choice studies description of procedures is too brief to permit comment on the influence of the instructions and of *E*'s behavior in the situation.

Mixture of Trinkets and Edibles

Trinkets and edibles may be used together to provide (a) variation in strengths of reinforcers and (b) variety among reinforcers to forestall satiation in experiments requiring many sessions. The procedure in a problem-solving study by Kendler, Kendler, Pliskoff, and D'Amato (20) is an example of the first type of usage. Interest was centered on the influence of reinforcement during training and during the test trials on inferential behavior. Some responses (pushing a lever inward) produced a raisin as a

CHILD DEVELOPMENT

weak reinforcer and a trinket or charm as a strong reinforcer. An example of the second usage of a mixture of trinkets and edibles is a methodological investigation conducted at this Laboratory in 1956. The question was whether a mixture would produce, over a number of sessions, stable response patterns which could serve as a baseline for evaluating other experimental variables. If it could be accomplished, the way would be opened to study systematically social, emotional, and symbolic conditions on individual subjects. The mixture was trinkets, dime-sized cookies, colored corn cereal (Trix), and M&M candy. Reinforcers were available on delivery according to one of four schedules—fixed ratio, variable ratio, fixed interval, and variable interval. The Ss, 18 in number, were from a private day-care center and ranged in age from 2-6 to 5-0. All were seen more than once; some as many as 10 times. There was a total of 81 sessions with intervals between sessions ranging from 2 to 17 days. In essence, it was found that the mixture strengthened lever-pressing behavior and maintained it at fairly constant rates in all but two children. One was a 4-year-old boy whose rate in the first session was abnormally slow compared to the other Ss. In subsequent sessions his rate gradually decreased, and finally he refused to return. His history and his behavior in the nursery school suggested severe emotional problems. The other exception was a 2½-year-old child—the youngest in the group. He refused to play with the trinkets or eat the edibles. Everything that came from the chute was pushed in the hole in the lever box. His rate of responding showed negligible change following reinforcement, and he spent more and more time playing with the other toys and pulling at the window shades. These two exceptions are further examples of the influence of age and historical factors on the effectiveness of experimentally manipulated reinforcers.

Two years later this procedure was used in a single-lever-type discrimination study with the following refinements: (a) the mixture was trinkets, dime-sized cookies, corn cereal (Sugar Corn Pops), and specially prepared sugarless candy in the proportions of 5 : 3 : 3 : 1; (b) each session was 20 minutes long; and (c) intervals between sessions were six days. Complicated discriminations were learned (two minutes of responding on a variable ratio of 25 or 50 on the occasion of an amber light and two minutes of no responding on a blue light) and some of the Ss were seen as many as 12 times, with maintenance of stable performance.

Hobby Items

Hobby items should serve satisfactorily as reinforcers for middle childhood, preadolescent, and adolescent Ss. The items would, of course, have to be selected in accordance with the known hobbies of the Ss. One of two procedures could be followed—tailor the items to each S's hobbies or select a group of Ss with the same hobby and use a common set of reinforcers. Many possibilities for suitable material come to mind—stamps, small jewelry, items such as pop beads, pictures of baseball players, airplanes, boats, historical figures, and the like. Unfortunately, examples of research

using hobby items cannot be given since none has as yet been reported in the literature.

Summary

The use of toys, trinkets, trinkets and edibles, and hobby items as reinforcers in experiments with children has been considered. Toys may be used in multiple-choice types of experiments to give a child a limited play period during the experiment or given as a gift during or at the end of an experiment. Studies have been described using these procedures or some combination of them. It was pointed out that some of the practices cited involved the active participation of the experimenter, making it difficult to separate the reinforcing property of the toy from the behavior of the experimenter. Trinkets, on the other hand, are usually given as gifts. Sometimes they are displayed following a response and given at the end of the experimental period, and sometimes they are immediately available following a response. The first procedure was illustrated by two studies on preschool children, one using a restricted operant, the other a free operant. The second procedure was exemplified by a free operant study on preschool children, and two multiple-choice situations, one on middle childhood children and one on the retarded. A combination of trinkets and edibles may be used to introduce variation in strength of reinforcers or to supply variety in experiments requiring a large number of sessions. One multiple-choice study on preschoolers was cited to illustrate the first, and two free operant investigations on preschoolers were cited to illustrate the second. Finally, it was pointed out that hobby items can serve as reinforcers for children above the preschool level. Procedures might involve tailoring the item to each subject's interest or preselecting subjects with a common interest and using a common set of hobby items.

The following suggestions for using manipulables as reinforcers seem indicated:

1. In the suggestions for using consumables as reinforcers, it was recommended that prior to initiating a study, parents, teachers, principals and/or superintendents be told what, in general, the experiment is about and what kind of reinforcers would be involved. Naturally, the same practice is recommended when manipulables are to be used. If it is planned to give manipulables to the child to keep, it may be reassuring to show some samples. In the event of objections to children's receiving toys, trinkets, hobby items, and perhaps edibles as well, it is suggested that the experimental procedure be modified to use limited toy play in situations, as in the Kendler, Kendler, Pliskoff, and D'Amato study (20). If this type of reinforcement is not appropriate for the study, it is suggested that motion pictures, TV, music, or social reinforcers be tried. [See articles by Jeffrey (17) on music and by Gewirtz and Baer (14) on social reinforcers.]

2. In selecting manipulables for a study, it will be advantageous to take several things into account. First, they ought to be appropriate for the particular group of subjects under study. That is, the material ought to

CHILD DEVELOPMENT

be chosen in the light of such factors as the age, sex, socioeconomic status, and living conditions (residential institution, day care centers, private school, etc.). Second, materials that may readily cause hurt or disturbances in the school or home should be avoided. These would include buttons that can easily be put into the mouth or nose, noise makers, wind-up toys, balls, and symbolic objects that may be contrary to parents' practices and teachings, e.g., certain types of religious symbols or items suggesting aggression. Third, for the convenience of interested investigators the materials used should be easily identifiable and obtainable. And fourth, it is recommended that mixtures of manipulables or mixtures of manipulables and edibles be considered in studies requiring many repeated measures.

3. It is suggested here again that reinforcers other than those experimentally introduced be maximally controlled by using minimum amount of instructions (to bring the child to the situation and to explain the task) and delivered in a consistent and standard fashion, and by presenting the manipulables in such a way that social factors are eliminated, controlled, or evaluated.

4. To make clear the conditions under which the reported behavior was observed, it is suggested that complete and detailed information be given in the procedure section of the report of the study. That portion dealing with reinforcement should tell what type of manipulable was received, and how it was received. If given during the session, what was the child allowed to do with it? If displayed during the session, how was it displayed and what were the instructions prior to the experiment? If delivered by E, what was E allowed to do and say at the time of the delivery and between deliveries?

REFERENCES

1. ALBERTS, E., & EHRENFREUND, D. Transposition of children as a function of age. *J. exp. Psychol.*, 1951, 41, 30-38.
2. AZRIN, N. H., & LINDSLEY, O. R. The reinforcement of cooperation between children. *J. abnorm. soc. Psychol.*, 1956, 52, 100-102.
3. BAER, D. M. A use of the free operant in the study of children's behavior. *Stud. Res. Psychol., Univer. Chicago*, 1957, 1, 5-18.
4. BIJOU, S. W. A systematic approach to an experimental analysis of young children. *Child Developm.*, 1955, 26, 161-168.
5. BIJOU, S. W. Patterns of reinforcement and resistance to extinction in young children. *Child Developm.*, 1957, 28, 47-54.
6. BIJOU, S. W. Methodology for an experimental analysis of child behavior. *Psychol. Rep.*, 1957, 3, 243-250.
7. BIJOU, S. W. Operant extinction after fixed interval schedules with young children. *J. exp. Anal. Behav.*, 1958, 1, 25-29.

SIDNEY W. BIJOU and PERSIS T. STURGES

8. BRACKBILL, YVONNE, & O'HARA, J. Discrimination learning in children as a function of reward and punishment. Paper read at Western Psychol. Ass., Eugene, Oregon, 1957.
9. CALVIN, A. D., & CLIFFORD, L. T. The relative efficiency of various types of stimulus-objects in discriminative learning by children. *Amer. J. Psychol.*, 1956, 69, 103-106.
10. CALVIN, A. D., PERKINS, M. J., & HOFFMAN, F. K. The effect of non-differentiated reward and non-reward on discrimination learning in children. *Child Developm.*, 1956, 27, 439-446.
11. CANTOR, G. N., & HOTTEL, J. V. Discrimination learning in mental defectives as a function of magnitude of food reward and intelligence level. *Amer. J. ment. Def.*, 1955, 60, 380-384.
12. FERSTER, C. B. The use of the free operant in the analysis of behavior. *Psychol. Bull.*, 1953, 50, 263-274.
13. FULLER, P. R. Operant conditioning of a vegetative human organism. *Amer. J. Psychol.*, 1949, 62, 587-590.
14. GEWIRTZ, J. L., & BAER, D. M. The effect of brief social deprivation on behaviors for a social reinforcer. *J. abnorm. soc. Psychol.*, 1958, 56, 49-56.
15. HOUSE, BETTY J., ZEAMAN, D., ORLANDO, R., & FISCHER, W. Learning and transfer in mental defectives. Progress Report, No. 1, Oct., 1957.
16. HUNTER, I. M. L. An experimental investigation of the absolute and relative theories of transpositional behavior in children. *Brit. J. Psychol.*, 1952, 43, 113-128.
17. JEFFREY, W. E. New techniques for motivating and reinforcing children. *Science*, 1955, 121, 371.
18. KELLER, F. S., & SCHOENFELD, W. N. *Principles of psychology*. New York: Appleton-Century-Crofts, 1950.
19. KENDLER, H. H., & KENDLER, TRACY S. Inferential behavior in pre-school children. *J. exp. Psychol.*, 1956, 51, 311-314.
20. KENDLER, H. H., KENDLER, TRACY S., PLISKOFF, S. S., & D'AMATO, MAY F. Inferential behavior in children: 1. The influence of reinforcement and incentive motivation. *J. exp. Psychol.*, 1958, 55, 207-212.
21. KOGAN, C. L. Withdrawal of a positive reinforcer and strength of behavior in children. Unpublished doctoral dissertation, Univer. of Washington, 1956.
22. KUENNE, MARGARET R. Experimental investigation of the relation of language to transposition behavior in young children. *J. exp. Psychol.*, 1946, 36, 471-490.

CHILD DEVELOPMENT

23. LASKO, A. A. A theoretical study of partial reinforcement within the framework of Rotter's social learning theory of personality. Unpublished master's thesis, Ohio State Univer., 1950.
24. LEWIS, D. J. Partial reinforcement in a gambling situation. *J. exp. Psychol.*, 1952, 43, 447-450.
25. LINDSLEY, O. R. Status Report III, Report of the first half of the second year on the Study of Psychotic Behavior, Studies in Behavior Therapy, Metropolitan State Hospital, Waltham, Mass., 1954.
26. NOWLIS, V., & NOWLIS, HELEN H. Child psychology. In *Ann. Rev. Psychol.*, 1952, 3, 1-28.
27. PERKINS, M. J., BANKS, H. P., & CALVIN, A. D. The effect of delay on simultaneous and successive discrimination in children. *J. exp. Psychol.*, 1954, 48, 416-418.
28. PLENDERLEITH, MAVIS. Discrimination learning and discrimination reversal learning in normal and feeble-minded children. *J. genet. Psychol.*, 1956, 88, 107-112.
29. PUMROY, D. K. The effects of intermittent reinforcement on resistance to extinction and emotional behavior in preschool children. Unpublished doctoral dissertation, Univer. of Washington, 1954.
30. PUMROY, SHIRLEY A. S. The effects of amount of reinforcement on extinction and emotional behavior with preschool children. Unpublished doctoral dissertation, Univer. of Washington, 1954.
31. ROBERTS, K. E. The ability of preschool children to solve problems in which a simple principle of relationship is kept constant. *J. genet. Psychol.*, 1932, 40, 118-135.
32. SARASON, S. B. *The clinical interaction*. New York: Harper, 1954.
33. SIEGEL, P. S., & FOSHEE, J. G. The law of primary reinforcement in children. *J. exp. Psychol.*, 1953, 45, 12-14.
34. SKINNER, B. F. *Science and human behavior*. New York: Macmillan, 1953.
35. STEVENSON, H. W. Latent learning in children. *J. exp. Psychol.*, 1954, 47, 17-21.
36. VERPLANCK, W. S. *A glossary of some terms used in the objective science of behavior*. Washington, D.C.: Amer. Psychol. Ass., 1957.
37. WARREN, A. B., & BROWN, R. H. Conditioned operant response phenomenon in children. *J. gen. Psychol.*, 1943, 38, 181-207.

ADVICE SEEKING OF MOTHERS AS A FUNCTION OF NEED FOR COGNITION

J. STACY ADAMS*¹

Stanford University

Casual observation of human behavior suggests strongly that some individuals seek organization, stability, and structure more than others. Theoretical suggestions to the same effect have also been made (3, 4, 6) and a need for cognition has been postulated. Cohen (3) has defined need for cognition as "a need to structure relevant situations in meaningful, integrated ways . . . to understand and make reasonable the experiential world" (p. 291). It is assumed that frustration of this need, as when understanding is not possible in an ambiguous situation, results in tension and deprivation. As a result, behavior will be directed toward structuring the ambiguous situation and increasing understanding. Cohen and his associates (1, 3) have provided experimental evidence for the existence of need for cognition and have developed an instrument for its measurement. As operationally measured, the need appears independent of need achievement (3), and data from the present study suggest that it is independent of education.

Given the existence of a measurable need for cognition, as defined, it may be hypothesized that a mother whose child presents a behavioral problem which she cannot handle or about which she is uncertain will experience tension and will try to reduce the tension if she has a need for cognition. One way of reducing the attendant tension under these circumstances is to seek information or advice about the problem from an appropriate source. An *appropriate* source is seen as one which, in the past, under similar circumstances has reduced tensions resulting from frustration of the need for cognition. This follows from the theoretical consideration that any stimulus (e.g., a source) which has been repeatedly associated with the reduction of tension (e.g., that resulting from the frustration of the need for cognition)

* Department of Psychology, Stanford University, Stanford, California.

¹ The research reported herein was supported in whole by Public Health Service grant 3M-9122(C1), United States Department of Health, Education, and Welfare, to whom the author is indebted.

CHILD DEVELOPMENT

will tend to elicit that response which preceded tension reduction (e.g., seeking information) (e.g., 5). A general hypothesis, then, is that *mothers with a comparatively high need for cognition who are faced with a behavioral problem in their child will tend to seek information or advice from a source in the presence of which, in the past, tension from a frustrated need for cognition arising under similar circumstances has been reduced.*

In order to predict from *what* sources information or advice will be sought about child behavior problems, it is necessary to have prior information on the frequency with which various sources have reduced tensions under similar circumstances. Unfortunately, precise data on this point are not readily available. However, some evidence regarding the relative frequency with which some sources are consulted is available. Since the frequency of tension reduction is very likely related to the frequency of consultation, the latter may be used as an index of the former. Considering physicians, psychological counselors, and school counselors, it can be said that physicians are seen more frequently than the other two classes of consultants and that, therefore, they will have more frequently reduced their advisees' tensions. It is specifically hypothesized, as a consequence of these considerations, that *mothers with a high need for cognition will more frequently seek information or advice about behavior problems in their child from physicians than will mothers with low need for cognition, whereas there will be no difference between mothers with high and low need for cognition in the frequency with which they seek advice from psychological and school counselors.*

The operational test of the hypothesis will be to determine whether mothers who score high on Cohen's measure of need for cognition (3) are more likely than mothers who score low to seek advice from physicians, while no such differentiation can be made with regard to seeking advice from psychological and school counselors.

METHOD

Ninety-five mothers with at least one child in the first grade of the nine public schools of Sunnyvale, California, were personally interviewed by trained survey interviewers on two occasions, two weeks apart. Selection of mothers was accomplished by selecting the name of every eleventh parent on the first grade school rosters. During the first interview 100 mothers were asked a series of questions bearing on the sources of advice they consulted when confronted with a behavioral problem in one of their children, on personal information, and on a number of other subjects not pertaining to the present study. Of the 100 mothers, 95 were reinterviewed two weeks later and were asked to answer the short version of Cohen's Situations Questionnaire, as well as a few other questions. The short version contains the 10 items that have the characteristic that responses to them intercorrelate to a higher degree than responses to any of the other items on the original Situations Questionnaire.

J. STACY ADAMS

All the interviews were conducted in the mothers' homes at their convenience, although usually not by appointment. The first interview was introduced as being conducted by Stanford University and as being concerned with "what parents think." The second was introduced as a follow-up. All questions asked were pretested on independent samples prior to the study, except for items on the Situations Questionnaire, which were considered as having been pretested by Cohen and his associates (3). The Situations Questionnaire was administered as by Cohen except for slight changes in instructions to respondents and the substitution of feminine for masculine personal pronouns. None of the respondents appeared to have had difficulty with the instruments.

RESULTS

The answers to three questions asked during the course of the first interview and the scores made on the Situations Questionnaire administered during the second interview furnish the data appropriate to a test of the hypothesis. Table 1 shows the relationship between need for cognition and the seeking of advice about child problems from physicians. The data clearly show that mothers with a high degree of need for cognition are more likely than mothers with a low need to seek advice. Thus, whereas 62 per cent of those with the highest need had sought advice from a doctor about child problems other than health, only 20 per cent of those with the lowest need had done so. The difference in distribution of advice-seeking for the four categories of need for cognition are significant beyond the .025 level.

In further support of the hypothesis and in contrast to the results already cited, Tables 2 and 3 show that mothers with high need for cognition are not more likely than those with low need to seek advice from psychological and school counselors. The hypothesis under test in this study is therefore fully confirmed.

TABLE 1
INCIDENCE OF ADVICE-SEEKING FROM PHYSICIANS BY AMOUNT
OF NEED FOR COGNITION

Advice Sought from Physician*	NEED FOR COGNITION SCORE							
	0-2		3		4		5-10	
	%	N	%	N	%	N	%	N
Yes	20	4	39	11	57	13	62	15
No	80	16	61	17	43	10	38	9
Total	100	20	100	28	100	23	100	24
$\chi^2 = 9.61, 3 \text{ df}, p < .025$								

* The question was: "Have you ever asked advice from a doctor about any problem with your child(ren), other than health problems?"

CHILD DEVELOPMENT

TABLE 2

INCIDENCE OF ADVICE-SEEKING FROM PSYCHOLOGICAL COUNSELORS BY AMOUNT OF NEED FOR COGNITION

<i>Advice Sought from Psychological Counselor*</i>	NEED FOR COGNITION SCORE							
	0-2		3		4		5-10	
	%	N	%	N	%	N	%	N
Yes	10	2	11	3	—	—	—	—
No	90	18	89	25	100	23	100	24
Total	100	20	100	28	100	23	100	24

$$\chi^2 = 5.16, 3 \text{ df}, .20 > p > .10 \text{ (uncorrected)}$$

* The question was: "Have you ever asked advice from a psychological counselor about any problem with your child(ren)?"

TABLE 3

INCIDENCE OF ADVICE-SEEKING FROM SCHOOL COUNSELORS BY AMOUNT OF NEED FOR COGNITION

<i>Advice Sought from School Counselor*</i>	NEED FOR COGNITION SCORE							
	0-2		3		4		5-10	
	%	N	%	N	%	N	%	N
Yes	—	—	4	1	4	1	9	2
No	100	20	96	27	96	22	91	21
Total	100	20	100	28	100	23	100	23†

$$\chi^2 = 2.03, 3 \text{ df}, .70 > p > .50 \text{ (uncorrected)}$$

* The question was: "Have you ever asked advice from a school counselor about any problem with your child(ren)?"

† One respondent did not answer above question.

DISCUSSION

If the assumption is correct that physicians have more frequently reduced tensions that are similar to those produced by the frustration of a need for cognition in the area of child behavior problems than have psychological and school counselors, the results of the present study have supported the hypothesis that mothers with high need for cognition will be more likely to seek advice about child behavior problems from physicians than from counselors. This seems altogether "understandable." Most individuals have had, during the course of their lives, some contact with physicians in connection with somatic complaints and, perhaps, psychosomatic complaints and "adjustment" problems. It is reasonable to suppose that on frequent occa-

J. STACY ADAMS

sions, if not invariably, the tensions and anxieties associated with the complaints and problems have been reduced. On the other hand, individuals infrequently come in contact with psychological and school counselors; as a result, the opportunities for tension and anxiety reduction are small. It follows, therefore, that, when tensions and anxieties arise under circumstances similar to those that led previously to their reduction, individuals will more likely visit a physician than a psychological or school counselor. In the present instance, tensions resulting from behavioral problems in one's own children are seen as being "similar" to those arising from somatic, psychosomatic, and psychological complaints in oneself or to those arising from past problems with children.

It might be suggested, of course, that mothers go to physicians rather than to counselors for advice because they are not aware of the existence or availability of the latter. This does not account for the *differential* likelihood of visiting physicians by mothers with high and low need for cognition, however. Another explanation of the results might be as follows: The results are an artifact because so few mothers visited psychological and school counselors that there was no opportunity to demonstrate a significant relationship with need for cognition. Such an explanation is fallacious, however, for given the marginals of Tables 2 and 3 a significant association between advice seeking and need for cognition *can* be obtained—for example, if the five cases seeking advice from a psychological counselor and the four cases seeking advice from a school counselor had also had the highest need for cognition scores.

SUMMARY

It was hypothesized that mothers with a comparatively high need for cognition would more frequently seek information or advice about behavior problems in their child from physicians than would mothers with low need for cognition, whereas no such differentiation would exist with respect to advice seeking from psychological and school counselors.

The test of the hypothesis was conducted by interview with 95 mothers who had at least one child in the first grade. The hypothesis was fully confirmed. The findings were discussed in the context of drive reduction theory.

REFERENCES

1. COHEN, A. R. Need for cognition and order of communication as determinants of opinion change. Ch. 6 in C. I. Hovland *et al.*, *Order of presentation in persuasion*. New Haven: Yale Univer. Press, 1957.
2. COHEN, A. R. Need for cognition, cognitive tuning and the organization of contradictory information. Unpublished manuscript, February, 1958.

CHILD DEVELOPMENT

3. COHEN, A. R., STOTLAND, E., & WOLFE, D. M. An experimental investigation of need for cognition. *J. abnorm. soc. Psychol.*, 1955, 51, 291-294.
4. FRENKEL-BRUNSWICK, ELSE. Intolerance of ambiguity as an emotional and perceptual personality variable. *J. Pers.*, 1949, 18, 108-143.
5. HULL, C. L. *Principles of behavior*. New York: Appleton-Century, 1943.
6. MASLOW, A. H. A theory of human motivation. *Psychol. Rev.*, 1943, 50, 370-396.

SOME BODY MEASUREMENTS OF NATIVE-BORN WHITE CHILDREN OF SEVEN TO FOURTEEN YEARS IN DIFFERENT CLIMATIC REGIONS OF TEXAS¹

JESSIE WHITACRE* and ETHEL T. GRIMES*

Agricultural and Mechanical College of Texas

In 1937 and 1938 the Texas Agricultural Experiment station along with participants in 13 other states and the District of Columbia cooperated in a study of body measurements for sizing children's garments and patterns. This study was initiated and led by the Bureau of Home Economics of the United States Department of Agriculture. Thirty-six measurements were made on each of 3000 boys and 3535 girls of 7 to 14 years of age in the public schools in four places of the state. No other study in Texas had included so many children or accumulated so many dependable data on each child. Therefore, by selection of appropriate measurements, nutrition workers were offered a unique opportunity to study body size of Texas children in relation to locality and to occupational groups. Such factors within states were not considered by the Bureau in analyzing the total mass of data (8, 9). Analysis of the Texas data was begun in 1940, then by unforeseen circumstances held in abeyance, but recently completed. This report is presented as a contribution to fundamental information for those concerned with problems of children's growth.

PROCEDURE

San Antonio, Houston, Lubbock and Denton (Figure 1) were selected as the places for this study. Denton included the relatively small number of records from nearby Dallas. In 1936 the population of Houston was 292,352, Dallas 260,475, San Antonio 231,542, Lubbock 20,520, and Denton 9,587. Local public school authorities selected schools where enrollment

* Department of Home Economics, Agricultural and Mechanical College of Texas, College Station.

¹ Published with the approval of the director of the Texas Agricultural Experiment Station as technical article No. 2900.

CHILD DEVELOPMENT

was predominantly Anglo-American and the pupils were from families which, by socioeconomic and intellectual considerations, they regarded as representative of that locality. All records used were of native-born white children only, both of whose parents also were born in the United States. This procedure constituted the most feasible approach to racial homogeneity and assured at least second-generation citizenship of the children measured. Among Texas 9-year-old children, for example, approximately two-thirds of the fathers and three-fourths of the mothers were born in Texas, and the remaining number of the parents in other states.

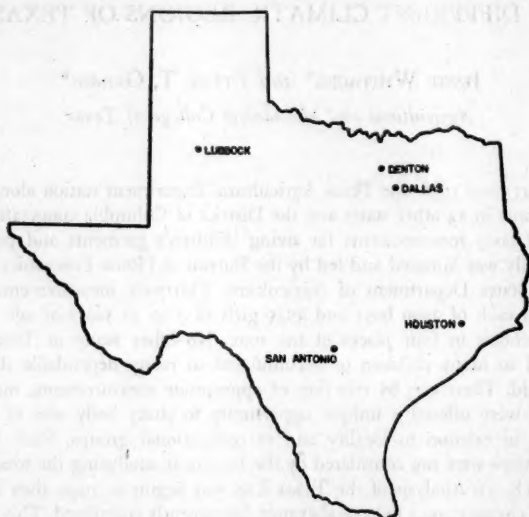


FIGURE 1—Location of places in which children were measured.

The measuring of the children followed the standardized procedures, details of which are described in USDA Miscellaneous Publication 365 (8) and illustrated and discussed in USDA Miscellaneous Publication 366 (9). Measurements, at the time they were made for each child, were recorded on the printed schedule form especially designed for this purpose. This form also carried supplementary information as to the date of measurement, the child's sex, name, address, whether rural or urban residence, place and date of birth, the birthplace of each parent, and the occupation of the principal wage earner, or supporter, of the child's family. The occupation of the family's supporter was the chief criterion of socioeconomic classification in either the "A" group (professional and skilled workers), or the

"B" group (semiprofessional or unskilled workers, or if ever recipient of relief). Before the original records were forwarded to the Bureau, each was transcribed onto a duplicate mimeographed form and, with the Bureau's consent, retained for use in this study.

The data for each child—selected measurements, place, economic and occupational groups, sex, and age to nearest birthday—were transferred to I.B.M. cards. The consistent results from preliminary analyses for 9-year-old children using 19 measurements indicated that fewer measurements would be satisfactory. For this report the greatest number used is 11, including: weight, bitrochanteric diameter, stature, height of waist, and of tiabiale, girth of arm, chest, waist, hip, thigh, and calf. Further reduction to seven, five, and four measurements sufficed for some purposes, but in each instance the different types of dimensions were represented—length, width, girth, bulk.

Successively the data were subjected to analyses for all ages combined, and for single year-of-age groups within place; all ages combined for each of six occupational groups within place; single year-of-age within occupational groups, places combined; and stature-hip girth classes within place. Supplementary consideration was given to the climatic characteristics of the four places as indicated by the experiment station weather records for 1923-1937 inclusive, the years covering the children's growth period. The statistical treatments employed included the calculation of means of measurement with standard error, standard deviation, coefficient of variability, and analysis of variance.

RESULTS AND DISCUSSION²

Distribution of Children

Table 1 shows for boys and girls, in each of the four places, the total number for all ages and the percentage distribution among the age groups. The number per age in Denton approximated one-third to one-half the number in the corresponding age in the other three places. The percentage distribution was similar, on the whole, in the four places, although Lubbock and Denton had a somewhat smaller proportion in the 13- and 14-year-old groups and larger in the 7-year group for both boys and girls than San Antonio and Houston had. Also, the proportion of 9- and 10-year-old boys

² The following supplementary material has been deposited as Document number 5835 with the ADI Auxiliary Publications Project, Photoduplication Service, Library of Congress, Washington 25, D. C.: (a) list of occupations within occupational classes; (b) means for each of seven measurements with standard errors and coefficients of variability for age groups within places; (c) means for each of four measurements with standard errors and coefficients of variability for occupational groups, places combined; (d) means for each of five measurements for hip girth classes within stature classes within places. A copy may be secured by citing the Document number and by remitting \$6.25 for photoprints, or \$2.50 for 35 mm. microfilm. Advance payment is required. Make checks or money orders payable to: Chief, Photoduplication Service, Library of Congress.

CHILD DEVELOPMENT

TABLE I
DISTRIBUTION OF CHILDREN BY AGE WITHIN PLACES

3 000 BOYS					3 535 GIRLS				
Year of Age	San Antonio	Houston	Lubbock	Denton	Year of Age	San Antonio	Houston	Lubbock	Denton
All ..	815	889	932	364	All ..	1001	1011	1096	427
PERCENTAGE OF CASES BY YEAR OF AGE									
7 ..	11.78	13.16	13.73	15.93	7 ..	11.59	11.77	10.86	16.39
8 ..	10.80	11.81	15.34	13.74	8 ..	10.09	15.03	14.05	11.94
9 ..	10.31	11.25	12.34	17.03	9 ..	12.19	12.17	15.24	11.94
10 ..	11.04	13.16	12.34	16.48	10 ..	8.69	11.87	9.58	12.65
11 ..	13.12	13.95	14.49	10.99	11 ..	11.29	11.97	16.51	14.99
12 ..	17.06	13.38	13.73	12.36	12 ..	15.68	12.96	14.87	13.12
13 ..	14.97	12.15	11.91	8.52	13 ..	15.88	12.56	10.95	11.01
14 ..	10.92	11.14	6.12	4.95	14 ..	14.59	11.67	7.94	7.96

and the 10-year-old girls was somewhat high in Denton, that of the 8-year-old boys and 9-year-old girls was a little high in Lubbock.

The distribution of boys and girls (ages combined) in six occupational groups are given in Table 2. The percentage distribution was strikingly similar for boys and girls in each occupational group. Percentage-wise, business and skilled labor were the predominating groups in all places. Lubbock

TABLE 2
DISTRIBUTION OF CHILDREN IN SIX OCCUPATIONAL GROUPS WITHIN PLACES, 7 TO 14 YEARS OF AGE INCLUSIVE

Places	Total No.	Prof'l	Semi- prof'l	Busi- ness	Skilled Labor	Common Labor	Farmer
Boys (N per group) ...	3000	281	294	1007	939	404	75
PERCENTAGE PER GROUP							
San Antonio	815	12.0	12.8	35.7	30.3	8.7	0.5
Houston	889	7.9	12.0	30.5	37.9	11.5	0.2
Lubbock	932	5.9	4.8	34.0	29.3	19.1	6.9
Denton	364	15.9	10.4	35.2	22.5	14.6	1.4
Grand total	3000	9.3	9.8	33.6	31.3	13.5	2.5
Girls (N per group) ...	3535	305	368	1222	1069	468	103
PERCENTAGE PER GROUP							
San Antonio	1001	9.9	13.0	38.2	29.7	8.6	0.6
Houston	1011	8.6	14.4	31.3	34.6	10.4	0.7
Lubbock	1096	5.8	5.4	34.3	30.3	16.9	7.3
Denton	427	12.9	7.7	34.7	20.9	21.5	2.3
Grand total	3535	8.6	10.4	34.6	30.3	13.2	2.9

JESSIE WHITACRE and ETHEL T. GRIMES

and Houston were relatively low in the professional group and Houston highest of the four places in skilled labor. Lubbock was lower than other places in the semiprofessional group, highest in farmer, and along with Denton, higher in common labor than other places. Only Lubbock had a sizable group of farmers. Notwithstanding these differences, the similarity in representation of the occupational groups (except farmer) is noteworthy.

The numbers in all groups except business and skilled labor in San Antonio, Houston, and Lubbock were too small for division into age groups. This fact together with the likeness in the percentage distribution in the four places led to combining places to make a distribution of occupational groups according to year-of-age for each sex (Table 3). The percentage distribution was fairly uniform among the occupational groups for each age, and for the age groups in separate occupational groups. Exceptions, 15 out of 96 values, were: 14-year-old boys whose proportion was low in all occupational groups, and 14-year-old girls low in professional and common labor; both boys' and girls' percentages at 13 years, low in the professional group; percentages of 8-year-old boys, and 7- and 10-year-old girls low in the farmer group, 7-year-old girls low in the semiprofessional group, and 10-year-old girls low in the skilled labor group. Still, it appears that the children represented comparable socioeconomic groups in the four places.

TABLE 3
DISTRIBUTION BY AGE WITHIN OCCUPATIONAL GROUPS,
INCLUDING ALL PLACES

<i>Year of Age</i>	<i>Prof'l</i>	<i>Semi-prof'l</i>	<i>Business</i>	<i>Skilled Labor</i>	<i>Common Labor</i>	<i>Farmer</i>
3000 Boys (<i>N</i> per group)	281	294	1007	939	404	75
PERCENTAGE OF CASES						
7	17.79	14.28	12.71	12.89	12.38	10.67
8	16.01	11.22	12.61	12.99	13.12	8.00
9	12.46	12.24	10.72	13.84	10.89	10.67
10	14.95	14.62	11.72	11.61	13.86	18.67
11	11.74	13.27	14.60	13.20	13.12	13.33
12	11.39	20.41	15.00	12.78	13.61	17.33
13	8.18	8.50	14.20	12.25	13.86	13.33
14	7.47	5.44	8.44	10.44	9.16	8.00
3535 Girls (<i>N</i> per group)	305	368	1222	1069	468	103
PERCENTAGE OF CASES						
7	13.11	8.42	10.96	14.50	11.97	7.77
8	14.75	13.04	13.01	12.63	12.39	12.62
9	14.43	12.50	12.77	12.91	14.10	12.62
10	10.16	11.68	11.29	9.17	11.11	3.88
11	18.03	13.59	13.18	11.51	15.38	17.48
12	13.11	17.12	15.38	13.47	11.32	18.45
13	7.87	10.60	12.52	14.22	15.17	13.59
14	8.52	13.04	10.88	11.60	8.55	13.59

CHILD DEVELOPMENT

Mean Measurements: Ages within Place

The first calculated means were for each of the 11 measurements with all ages combined, within each place. When the means were compared by ranking, San Antonio and Houston held first and second place respectively for 10 measurements, and their ranks were reversed for the one remaining measurement. Lubbock ranked 3 and Denton 4 in nine measurements for boys and eight measurements for girls, and ranks were reversed for the remaining measurements in each sex group. Boys and girls had the same ranking for each measurement with one exception—waist height for Lubbock rank 3 and for Denton rank 4 for boys, but the reverse for girls. Means for a given measurement of Lubbock and Denton children resembled each other more than either group resembled San Antonio or Houston.

The means for five of these measurements—weight, stature, hip girth, bitrochanteric diameter, calf girth—are shown graphically in Figures 2 through 6 for all ages together and for separate years of age in the four places. The separate year-of-age means in each of the five measurements present essentially the same picture as do the means for all years together; especially conspicuous is the excess of San Antonio and Houston means over those of Lubbock.

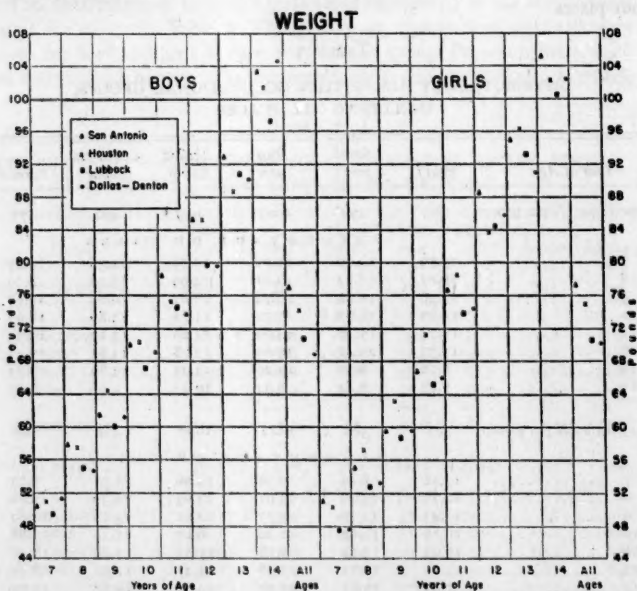


FIGURE 2—Mean weight of year-of-age groups, and all ages combined, of boys and girls in four places in Texas.

JESSIE WHITACRE and ETHEL T. GRIMES

Analysis of variance for these five measurements, all ages combined, showed that places were significantly different at the .01 level in each measurement (Tables 4 and 5). Denton was omitted in these analyses on account of small numbers.

To determine whether the lower proportion of 13- and 14-year-old children and the higher proportion of 7-year-olds in Lubbock and Denton accounted for the place differences for all ages combined, a variance analysis was made for each year-of-age group for boys and for girls. As shown in Tables 4 and 5, every year of age appeared among the significant differences, and 13- and 14-year-old groups appeared less frequently than did other ages. Significant differences were more numerous between San Antonio and Lubbock and between Houston and Lubbock than between San Antonio and Houston. Place differences were not accounted for by influence of age groups.

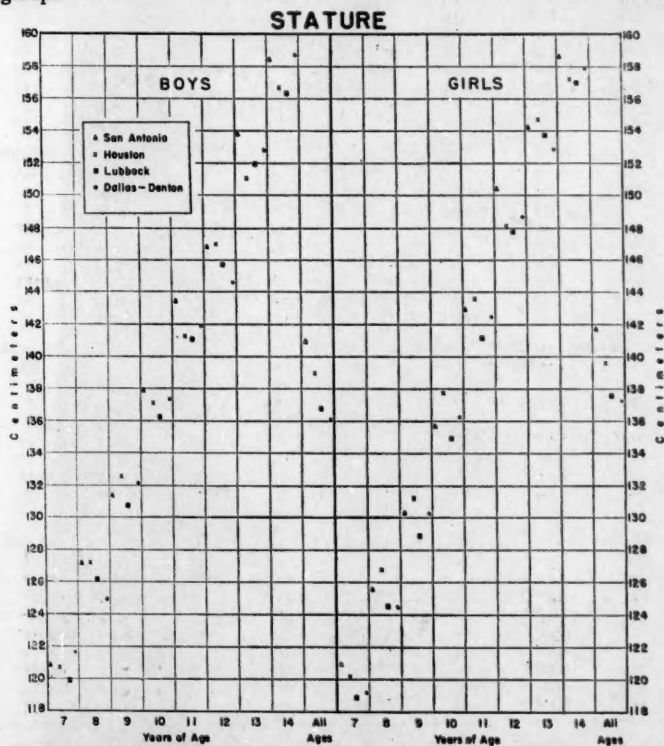


FIGURE 3—Mean stature of year-of-age groups, and all ages combined, of boys and girls in four places in Texas.

TABLE 4
PLACE DIFFERENCES IN MEAN MEASUREMENTS—BOYS

Age	San Antonio > Lubbock		San Antonio > Houston		Houston > Lubbock	
	F	Diff. t	Diff. t		Diff. t	
Weights (lb.)						
All ..	3.62**					
7 ..	< 1	0.97	—0.93		1.90	
8 ..	1.50	2.59	0.48		2.11	
9 ..	3.13*	1.37	—2.80		4.17	2.48*
10 ..	2.22	2.93	—0.15		3.08	
11 ..	3.43*	4.02	2.52*	3.19	1.96*	0.83
12 ..	8.86**	5.76	3.82**	0.34		5.42
13 ..	3.74*	2.09	4.45	2.73**	—2.36	3.45**
14 ..	5.32**	6.81	3.26**	2.41	4.40	2.15*
Stature (cm.)						
All ..	2.78**					
7 ..	< 1	1.00	0.19		0.81	
8 ..	1.56	1.19	—0.03		1.22	
9 ..	2.26	0.66	—1.13		1.79	2.11*
10 ..	1.98	1.72	1.97*	0.78	0.94	
11 ..	5.00**	2.32	2.89**	2.18	2.66**	0.14
12 ..	1.77	1.14	—0.22		1.36	
13 ..	5.91**	1.87	2.30*	2.74	3.34**	—0.87
14 ..	2.91	2.12	2.01*	1.89	2.09*	0.23
Hip Girth (cm.)						
All ..	3.25**					
7 ..	0.95	0.25	—0.65		0.90	
8 ..	1.78	1.25	0.36		0.89	
9 ..	6.09**	0.95	—1.54	1.99*	2.49	3.48**
10 ..	2.00	1.07	—0.23		1.30	
11 ..	2.95	1.64	2.42*	1.07	0.57	
12 ..	8.65**	2.27	3.54**	—0.18	2.45	3.67**
13 ..	1.13	0.50	1.04		—0.54	
14 ..	2.47	1.97	2.22*	0.61	1.36	
Biacromiatic (cm.)						
All ..	3.67**					
7 ..	< 1	0.13	—0.03		0.16	
8 ..	1.76	0.46	0.17		0.29	
9 ..	3.07*	0.33	—0.31		0.64	2.48*
10 ..	1.55	0.47	0.26		0.21	
11 ..	4.48*	0.73	2.98**	0.43	0.30	
12 ..	8.34**	0.81	3.50**	—0.04	0.85	3.53**
13 ..	7.04**	0.81	3.27**	0.79	0.02	
14 ..	2.82	0.75	2.34*	0.38	0.37	
Calf Girth (cm.)						
All ..	7.57**					
7 ..	7.41**	0.42	—0.63	2.15*	1.05	3.85**
8 ..	7.46**	1.02	0.23		0.79	2.89**
9 ..	12.41**	0.59	—0.86	2.73*	1.45	4.98**
10 ..	6.40**	0.74	2.47*	—0.22	0.96	3.43**
11 ..	4.16*	0.79	2.86**	0.34	0.45	
12 ..	16.32**	1.30	4.98**	—0.03	1.33	4.90**
13 ..	< 1	0.01	0.17		—0.16	
14 ..	6.16**	1.19	3.29**	0.12	1.07	3.02**

NOTE.—Significant F: all ages, 1% = 1.99, 5% = 1.64; single year of age, 1% = 4.60, 5% = 2.99. Significant t: single year of age, 1% = 2.58, 5% = 1.96.

* Significant at 5% level.

** Significant at 1% level.

TABLE 5
PLACE DIFFERENCES IN MEAN MEASUREMENTS—GIRLS

Age	San Antonio > Lubbock				San Antonio > Houston				Houston > Lubbock			
	F	Diff.	t		Diff.	t			Diff.	t		
Weight (lb.)												
All ..	4.33**											
7 ..	2.15	2.11			0.30				2.81			
8 ..	4.13*	1.95			-2.23				4.18	2.88**		
9 ..	4.27*	0.93			-3.37	2.07*			4.30	2.85**		
10 ..	7.88**	1.30			-5.07	2.83**			6.37	3.75**		
11 ..	5.20**	2.71			-2.03				4.74	3.18**		
12 ..	5.54**	3.90	2.74**		4.47	2.97**			-0.57			
13 ..	1.59	1.60			-1.29				2.89			
14 ..	3.90*	4.59	2.67**		2.89				1.70			
Stature (cm.)												
All ..	4.66**											
7 ..	3.84*	2.26	2.74**		0.88				1.38			
8 ..	5.33**	1.04			-1.32				2.36	3.27**		
9 ..	5.36**	1.46			-0.95				2.41	3.21**		
10 ..	5.84**	0.72			-2.05	2.31*			2.77	3.29**		
11 ..	5.43**	1.64	2.17*		-0.68				2.32	3.13**		
12 ..	8.53**	2.69	3.81**		2.42	3.24**			0.27			
13 ..	< 1	0.51			-0.47				0.98			
14 ..	2.18	1.55			1.33				0.22			
Hip Girth (cm.)												
All ..	4.89**											
7 ..	3.49*	1.95	2.39*		0.20				1.75	2.16*		
8 ..	5.30**	0.94			-1.38				2.32	3.25**		
9 ..	7.81**	1.18			-1.76	2.20*			2.94	3.96**		
10 ..	7.57**	1.10			-2.08	2.36*			3.18	3.81**		
11 ..	5.03**	1.43			-0.83				2.26	3.08**		
12 ..	6.17**	2.24	3.20**		2.09	2.82**			0.15			
13 ..	1.66	0.67			-0.78				1.45			
14 ..	2.11	1.74	2.05*		0.58				1.16			
Bitrochanteric (cm.)												
All ..	5.53**											
7 ..	3.94*	0.68	2.62**		0.11				0.57	2.21*		
8 ..	4.50*	0.31			-0.37				0.68	2.99**		
9 ..	2.39	0.14			-0.37				0.51	2.16*		
10 ..	6.52**	0.36			-0.59	2.10*			0.95	3.57**		
11 ..	9.73**	0.77	3.22**		-0.18				0.95	4.06**		
12 ..	8.30**	0.85	3.82**		0.74	3.14**			0.11			
13 ..	1.72	0.34			-0.11				0.45			
14 ..	7.17**	1.02	3.78**		0.40				0.62	2.20*		
Calf Girth (cm.)												
All ..	8.91**											
7 ..	6.27**	0.89	3.02**		-0.02				0.91	3.11**		
8 ..	10.22**	0.34			-0.80	2.76**			1.14	4.42**		
9 ..	9.37**	0.60	2.23*		-0.56				1.16	4.33**		
10 ..	15.13**	0.28			-1.27	4.00**			1.55	5.14**		
11 ..	13.92**	1.08	3.99**		-0.17				1.25	4.72**		
12 ..	6.61**	0.91	3.61**		0.56	2.10*			0.35			
13 ..	6.44**	0.72	2.64**		-0.27				0.99	3.45**		
14 ..	3.33*	0.76	2.49*		0.10				0.66	2.07*		

NOTE.—Significant F: all ages, 1% = 1.99, 5% = 1.64; single year of age, 1% = 4.60, 5% = 2.99. Significant t: single year of age, 1% = 2.58, 5% = 1.96.

* Significant at 5% level.

** Significant at 1% level.

CHILD DEVELOPMENT

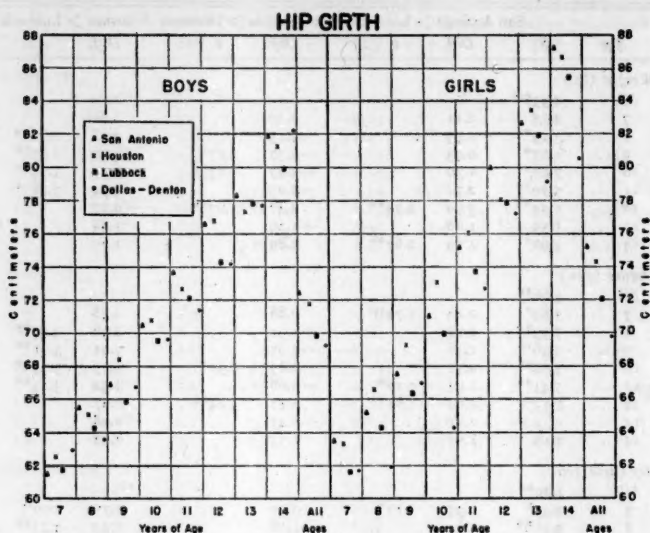


FIGURE 4—Mean hip girth of year-of-age groups, and all ages combined, of boys and girls in four places in Texas.

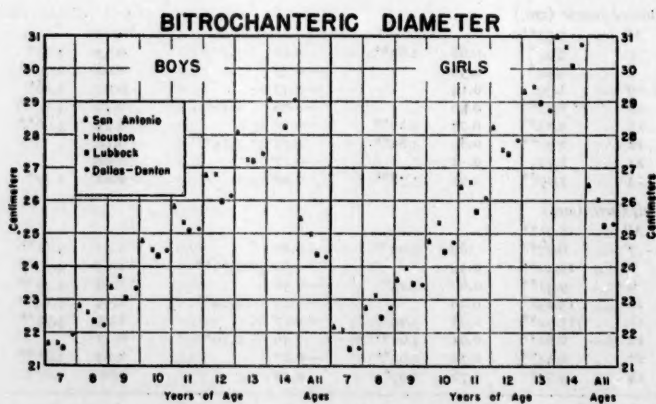


FIGURE 5—Mean bitrochanteric diameter of year-of-age groups, and all ages combined, of boys and girls in four places in Texas.

GIRTH OF CALF

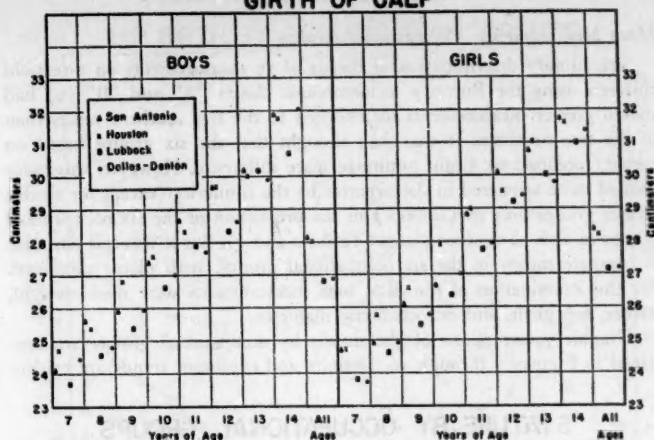


FIGURE 6—Mean calf girth of year-of-age groups, and all ages combined, of boys and girls in four places in Texas.

MEAN WEIGHT OF OCCUPATIONAL GROUPS

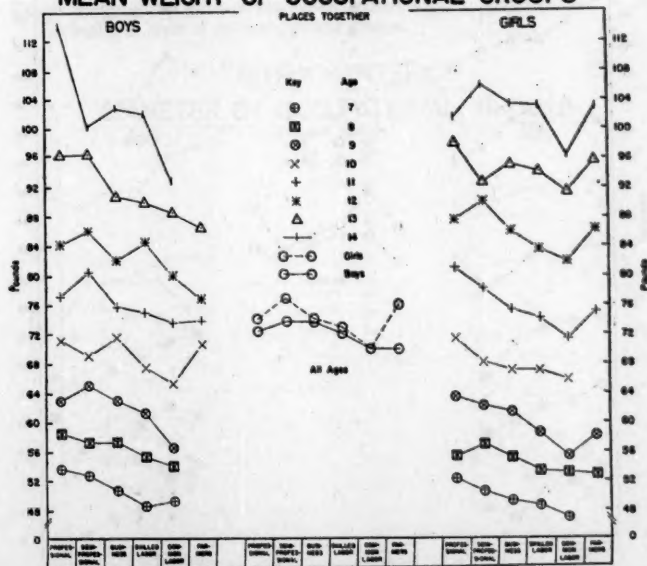


FIGURE 7—Mean weight of year-of-age groups, and all ages combined, of boys and girls in each of six occupational groups.

CHILD DEVELOPMENT

Mean Measurements: Occupational Groups

Preliminary determination of means of 19 measurements on 9-year-old children using the Bureau's socioeconomic classes "A" and "B" (9) had shown greater measurements for children in the two southern places than in the two northern. It was then thought that the six groups based on parents' occupations might minimize place difference. However, this point seemed to be answered in the negative by the similarity (except for sizable farmer groups only in Lubbock) in the proportion of the six occupational groups in each of the four places (Tables 2 and 3). But it was still desirable to compare means in the six occupational groups, with places combined. For this examination of the data, four measurements were used—weight, stature, hip girth, and bitrochanteric diameter.

Graphic presentations of the means by occupational groups are contained in Figures 7 through 10. Definite and consistent trends are evident

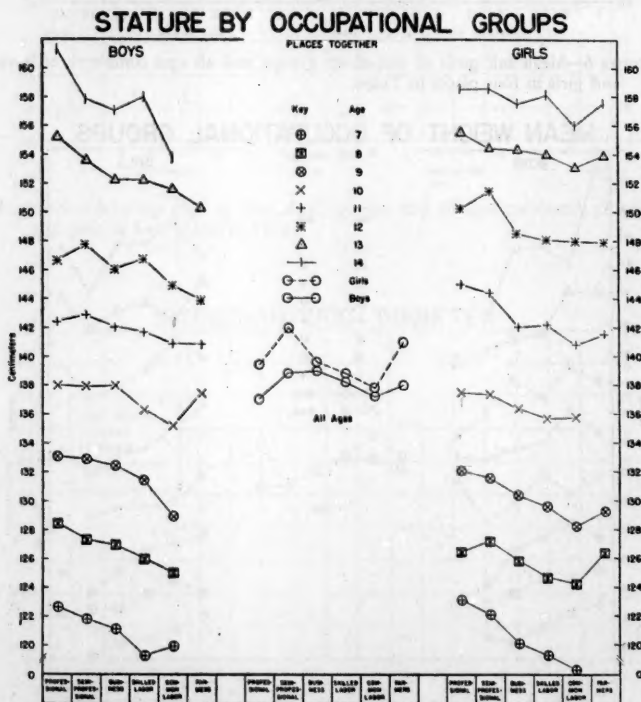


FIGURE 8—Mean stature of year-of-age groups, and all ages combined, of boys and girls in each of six occupational groups.

JESSIE WHITACRE and ETHEL T. GRIMES

HIP GIRTH BY OCCUPATIONAL GROUPS

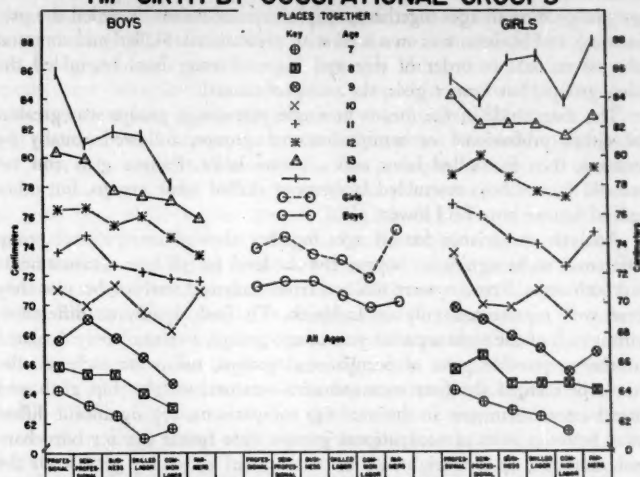


FIGURE 9—Mean hip girth of year-of-age groups, and all ages combined, of boys and girls in each of six occupational groups.

BITROCHANTERIC DIAMETER BY OCCUPATIONAL GROUPS

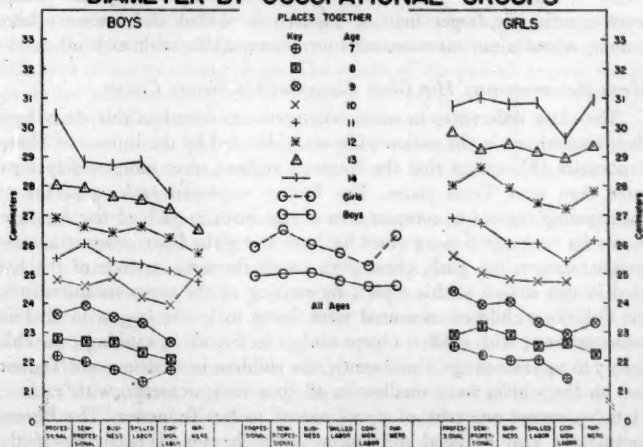


FIGURE 10—Mean bitrochanteric diameter of year-of-age groups, and all ages combined, of boys and girls in each of six occupational groups.

CHILD DEVELOPMENT

in all measurements for both the all-ages group and the separate year-of-age groups. For all ages together, semiprofessional means exceeded the professional, and business was on a level with professional. Skilled and common labor were next in order of size and mean. Farmer boys resembled the labor groups, but farmer girls, the semiprofessional.

The magnitude of the means in single year-of-age groups was greatest for either professional or semiprofessional groups, followed usually by business, then by skilled labor and common labor. Farmer girls and 10-year-old farmer boys resembled business or skilled labor groups, but other ages of farmer boys held lowest place.

Analysis of variance for all ages together showed occupational group differences to be significant beyond the .01 level for all four measurements and both sexes. Farmers were omitted from statistical analyses because they were well represented only in Lubbock. To find significant differences within each of the eight separate year-of-age groups, t values were calculated for the 10 possible pairs of occupational groups, using for each sex the means of each of the four measurements—stature, weight, hip girth and bitrochanteric diameter. In the total 640 comparisons, 177 significant differences between pairs of occupational groups were found (51 for bitrochanteric diameter, 47 for weight, 44 for stature, and 35 for hip girth). Of the total significant t values, only 25 were for the three pairs possible with professional, semiprofessional and business groups; only 10 for the skilled and common labor pairs; but 52 for skilled labor, and 90 for common labor, each compared with professional, semiprofessional, and business groups. The results of statistical analyses indicate that the mean measurements of children in professional, semiprofessional, and business groups were significantly larger than of children in skilled and common labor groups, whose mean measurements were comparable with each other.

Mean Measurements: Hip Girth Classes within Stature Classes

The place differences in mean measurements found in this study have their counterpart in the nation-wide study directed by the Bureau of Home Economics (8), except that the Bureau's regions were more widely separated than were Texas places. The Bureau employed ranking of the 15 participating regions to compare size of the mean in each of five measurements for each age 6 to 14 years for boys and girls. Four measurements—weight, stature, hip girth, chest girth—were the same as four of the five used in this section of this report. By ranking of the mean measurements, the California children measured were found to be the largest in all four measurements, with rank 1, except rank 2 in five of 64 rankings, for children 7 to 14 years of age. Consistently, the children in Alabama and Tennessee, on the whole, were smallest in all four measurements, with rank of 11 to 15, except one rank of 7 and one of 10 for Tennessee. The Bureau commented that regional differences might be found negligible if the children were classified on the basis of selected body measurements, disregarding age.

In line with the foregoing prediction, the children in each of the four places were classified by intervals of hip girth within stature intervals, these being the highly correlated body measurements recommended as the basis for sizing children's garments and patterns. Using this classification, the means were determined for each of five measurements—weight, bitrochanteric diameter, chest girth, calf girth, and waist height. Within each stature-hip girth class for each sex, the means within places were strikingly similar for each measurement. These similarities prompted calculation of maximum difference among either three or four places, each of which had 10 or more cases within a given stature - hip girth class. The maximum differences are shown in Tables 6 and 7 for boys and girls respectively.

One of the outstanding features of these tables is the great predominance of small size of maximum differences. With very few exceptions, these differences are inconsequential in comparison with the means in the stature-hip girth class from which they were derived. Maximum difference exceeded minimum mean within single stature - hip class, among the five measurements, by percentages of 0.3 to 6.2 for boys, 0.02 to 5.3 for girls. The maximum difference of the means for places, as percentage of the minimum mean, in a given stature - hip class, averaged for weight 2.5 and 1.8 for boys and girls, respectively; for bitrochanteric diameter, 1.5 and 1.2; chest girth, 1.6 and 1.9; calf girth, 2.3 and 2.3; and waist height, 1.2 and 1.3.

Another conspicuous feature of Tables 6 and 7 is the entire lack of trend among places with respect to the comparative size of maximum and minimum means for any of the five measurements, for either boys or girls. No one place had consistently a larger or smaller mean than any other place. Variance analyses were made to determine whether the means for places were statistically different, even though markedly similar in value, within stature - hip girth classes. Data used were the same as to calculate maximum difference of means among places. The results of the over-all analysis for all classes for each sex are given in Table 8. For boys, place differences were significant for means of four of the five measurements; only the bitrochanteric diameter mean, like that for the girls, was not significantly different between places. Girls had only two significant place differences, the mean for chest and calf girths.

Single stature - hip girth classes were also subjected to analysis of variance of means in places for each of the five measurements. As shown in Table 9, there were for boys only 15 stature - hip girth classes out of a total of 110 (22 classes \times five measurements) where means were significantly different among places; for girls there were 16 classes out of 125 (25 classes \times 5 measurements) with significant place differences. For both sexes, calf girth was the measurement with significant difference in the most classes (six for boys, eight for girls), followed by chest girth with five classes each for boys and girls. Although place difference for waist height of boys was significant at the .01 level and for weight at the .05 level in the over-all variance, only two single stature - hip girth classes showed significant place difference for each of these measurements (Tables 8 and 9).

CHILD DEVELOPMENT

TABLE 6
MAXIMUM DIFFERENCES BETWEEN PLACES IN MEAN MEASUREMENTS BY HIP GIRTH INTERVALS
WITHIN STATURE INTERVALS—BOYS

INTERVALS			MAXIMUM DIFFERENCE									
Stature	Hip Girth	No. of Places	Weight	Birotrochanteric Diameter	Client Girth	Calf Girth	Wrist Height					
cm.	cm.		lb.	cm.	cm.	cm.	cm.					
112.0-117.9 ..	57.0-60.9 ...	4	SA>L [†] ... 1.4	D>L 0.46	L>SA ... 1.37	H>L 0.71	SA>H ... 2.03					
118.0-123.9 ..	59.0-62.9 ...	4	L>D 1.1	SA>H ... 0.20	L>SA ... 0.79	H>L 0.48	SA>H ... 1.55					
	63.0-66.9 ...	3	H>SA ... 1.1	SA>H ... 0.18	H>SA ... 0.98	H>SA ... 0.84	SA>L ... 0.94					
124.0-129.9 ..	58.0-61.9 ...	4	H>L 1.5	D>SA ... 0.51	D>SA ... 1.31	SA>L ... 0.73	D>L ... 1.45					
	62.0-65.9 ...	4	D>L 2.0	D>H 0.23	D>H 0.49	D>L 0.68	SA>D ... 0.88					
	66.0-69.9 ...	4	H>D 1.7	D>H 0.39	L>D 1.30	SA>L ... 0.53	SA>D ... 0.46					
130.0-135.9 ..	59.0-63.9 ...	4	D>SA ... 2.8	D>SA ... 1.43	D>SA ... 2.41	D>L ... 1.08	D>SA ... 1.70					
	64.0-68.9 ...	4	SA<L, H>D ... 0.7	SA>H ... 0.07	L>SA ... 0.82	H>L ... 0.61	L>H ... 0.70					
	69.0-73.9 ...	4	L>SA ... 1.4	D>SA ... 0.44	L>SA ... 1.45	H>D ... 0.28	D>SA ... 1.33					
136.0-141.9 ..	62.0-66.9 ...	4	H>SA ... 1.4	D>L 0.56	L>SA ... 1.36	H>D ... 0.40	D>SA ... 0.68					
	67.0-71.9 ...	4	H>D>SA ... 0.6	SA>L ... 0.22	L>SA ... 0.64	H>L ... 0.59	SA>L ... 1.17					
	72.0-76.9 ...	4	SA>H ... 0.9	D>L 0.16	H>SA ... 1.40	H>L ... 0.75	SA>D ... 1.64					
142.0-147.9 ..	63.0-69.9 ...	3	SA>L ... 1.4	H>L 0.46	L>H 1.66	SA>L ... 1.40	L>H ... 0.85					
	70.0-74.9 ...	4	SA>L ... 1.4	SA>L ... 0.19	H>SA ... 0.81	SA>D ... 0.45	D>H ... 0.38					
	75.0-79.9 ...	4	SA>H ... 1.4	D>L 0.56	L>D ... 0.97	SA>L ... 0.60	D>H ... 1.28					
	80.0-84.9 ...	3	L>SA ... 5.7	L>SA ... 0.20	L>SA ... 4.30	L>SA ... 0.83	L>H ... 0.42					
148.0-153.9 ..	69.0-73.9 ...	3	SA>H ... 1.8	SA>L ... 0.39	L>SA ... 0.73	H>L ... 0.51	SA>L ... 1.24					
	74.0-78.9 ...	4	L>SA, H ... 1.1	D>SA ... 0.24	D>SA ... 1.05	H>SA ... 0.36	L>SA ... 0.48					
	79.0-83.9 ...	3	SA>L ... 1.6	L>SA ... 0.36	SA>L ... 0.43	SA>L ... 0.66	SA>H ... 0.88					
154.0-159.9 ..	73.0-77.9 ...	3	SA>H ... 2.9	SA>H ... 0.46	L>SA ... 0.66	SA>H ... 0.28	SA>H ... 1.20					
	78.0-82.9 ...	3	H>SA ... 1.2	H>L 0.42	H>L 0.92	H>L ... 0.56	H>SA ... 0.45					
160.0-165.9 ..	82.0-86.9 ...	3	L>H 4.0	SA>H ... 1.27	SA>H ... 1.37	L>H ... 0.34	H>L ... 0.44					

* With at least 10 cases in each place.

† Initial letters indicate places: SA, San Antonio; H, Houston; L, Lubbock; D, Denton.

TABLE 7
MAXIMUM DIFFERENCES BETWEEN PLACES IN MEAN MEASUREMENTS BY HIP GIRTH INTERVALS
WITHIN STATURE INTERVALS—GIRLS

INTERVALS		No. OF PLACES		MAXIMUM DIFFERENCES						
Stature	Hip Girth		Weight	Biochanteric Diameter	Chest Girth	Calf Girth	Wait Height			
cm.	cm.	lb.		cm.	cm.	cm.	cm.			
112.0-117.9 ..	58.0-61.9 ...	4	SA>D†	... 0.08	L>D	... 1.27	SA>D	... 0.74	SA>L	... 0.94
118.0-123.9 ..	56.0-59.9 ...	4	H>L	... 0.7	D>SA	... 1.19	H>SA	... 0.95	D>H	... 0.49
	60.0-63.9 ...	4	H>D	... 0.6	D>L	... 1.48	H>L	... 0.42	D>L	... 0.60
	64.0-67.9 ...	4	D>H	... 1.4	L>H	... 1.10	H>L	... 0.14	SA>L	... 1.09
124.0-129.9 ..	59.0-62.9 ...	4	H>L	... 1.0	D>SA	... 0.35	SA>D	... 0.50	H>L	... 0.61
	63.0-66.9 ...	4	SA>D, H	... 0.4	L>D	... 0.17	L>SA	... 0.97	H>L	... 0.81
	67.0-70.9 ...	4	D>SA	... 0.8	D>L	... 0.28	L>SA	... 1.16	SA>L	... 0.60
130.0-135.9 ..	61.0-65.9 ...	4	D>SA	... 0.6	L>D	... 0.30	L>SA	... 0.64	D>L	... 0.48
	66.0-70.9 ...	4	SA>L	... 0.7	D>H	... 0.35	L>SA	... 0.70	H>L	... 0.46
	71.0-75.9 ...	4	D>SA	... 1.5	H>L	... 0.36	L>SA	... 1.80	H>L	... 0.56
136.0-141.9 ..	65.0-69.9 ...	4	D>SA	... 0.9	SA>H	... 0.18	D>SA	... 1.22	D>L	... 0.59
	70.0-74.9 ...	4	D>H	... 0.9	D>L	... 0.38	D>SA	... 1.12	SA>L	... 0.46
	75.0-79.9 ...	3	H>L	... 0.3	SA>H	... 0.15	D>SA	... 1.80	H>SA	... 0.68
142.0-147.9 ..	69.0-73.9 ...	4	D>SA	... 0.9	D>H	... 0.42	L>SA	... 1.32	SA>L	... 1.15
	74.0-78.9 ...	4	SA>D	... 0.6	D>SA, L	... 0.31	H>SA	... 1.16	H>D	... 0.82
	79.0-83.9 ...	4	L>SA	... 3.1	H>L	... 0.40	L>SA	... 2.05	D>SA	... 0.63
148.0-153.9 ..	74.0-78.9 ...	4	D>L	... 1.7	D>SA	... 0.30	H>SA	... 1.23	H>L	... 0.63
	79.0-83.9 ...	4	H>D	... 3.3	SA>D	... 0.34	L>D	... 1.54	H>D	... 0.79
	84.0-88.9 ...	3	L>H	... 3.0	H>L	... 0.20	H>SA	... 0.45	H>L	... 0.04
154.0-159.9 ..	74.0-78.9 ...	3	SA>H	... 1.0	SA>L	... 0.48	H>SA	... 1.54	SA>L	... 0.91
	79.0-83.9 ...	4	D>H	... 1.7	D>L	... 0.34	L>SA	... 1.99	SA>L	... 0.70
	84.0-88.9 ...	3	SA>H	... 2.4	SA>L	... 0.48	L>SA, H	... 0.02	SA>H, L	... 0.33
	89.0-93.9 ...	3	L>SA	... 1.0	SA>L	... 0.42	H>SA	... 2.54	H>L	... 0.31
160.0-165.9 ..	83.0-87.9 ...	3	L>SA	... 1.7	SA>H	... 0.06	L>SA	... 1.60	H>SA	... 0.64
	88.0-92.9 ...	3	H>L	... 3.0	H>L	... 0.54	H>SA	... 1.45	H>SA	... 1.68

* With at least to cases in each place.

† Initial letters indicate places: SA, San Antonio; H, Houston; L, Lubbock; D, Denton.

CHILD DEVELOPMENT

TABLE 8

ANALYSIS OF VARIANCE OF MEASUREMENTS FOR ALL STATURE - HIP GIRTH CLASSES IN FOUR PLACES IN TEXAS

Measurement	B O Y S			G I R L S		
	MEAN SQUARES		F	MEAN SQUARES		F
	Between Places	Within Places		Between Places	Within Places	
Weight	18.00	13.00	1.38*	12.85	14.65	< 1
Waist height	7.11	4.16	1.71**	7.95	6.59	1.21
Chest girth	9.79	5.06	1.94**	12.77	6.27	2.04**
Calf girth	2.91	1.24	2.35**	2.94	1.52	1.93**
Bitrochanteric80	1.10	< 1	.74	.72	1.03

* Significant at 5% level ($F = 1.33$).** Significant at 1% level ($F = 1.48$).

Single class analyses for means of bitrochanteric diameter agreed with the over-all results—no significant difference for either sex group. Girls, without significant place difference for waist height in the over-all analysis, had significant differences in three single stature - hip girth classes.

The values in Table 10 confirm the conclusion drawn from Tables 6 and 7, viz., that place differences disappeared for the means in stature - hip girth classes. In each of the six pairs of places compared, the significant place differences (Table 10) within single stature - hip girth classes showed almost as many (—) values as (+) values. No one place had means that were consistently larger or smaller than any other place.

The foregoing examination of the data shows that children who, in any of the four places, had attained a given hip girth and stature, also had

TABLE 10

NUMBER OF SIGNIFICANT PLACE DIFFERENCES WITHIN SINGLE STATURE - HIP GIRTH CLASSES

Places* Compared	NO. OF DIFFERENCES			
	B O Y S		G I R L S	
	+	—	+	—
SA > L	8	5	8	5
SA > H	4	8	3	6
SA > D	2	3	2	2
H > L	10	3	9	4
H > D	4	4	3	3
L > D	2	4	4	3
Total no.	30	27	29	23

* Initial letters indicate places: SA, San Antonio; H, Houston; L, Lubbock; D, Denton.

TABLE 9

B O Y S							G I R L S						
C L A S S *		Weight	Birochanteric Diameter	Chest Girth	Calf Girth	Waist Height	C L A S S *		Weight	Birochanteric Diameter	Chest Girth	Calf Girth	Waist Height
Stature	Hip - Girth						Stature	Hip - Girth					
112	57				x	x	112	58					
118	59						118	56			x		
	63							60					
								64					
124	58				x		124	59			x	x	
	62							63					
	66							67					
130	59				x		130	61			x		
	64							66					
	60							71					
136	62						136	65					
	67							70					
	72							75					
142	65				x		142	69			x	x	x
	70							74					
	75							79					
	80							79					
148	69						148	74					
	74							79		x			
	79							84					
154	73						154	74			x	x	
	78							79		x			
								84					
								89					
160	82						160	83				x	x
								88					
Total no. classes							Total no. classes						
2							0						
5							5						
6							8						
2							3						

• The lower limit of the interval of each stature and hip-girth class.

CHILD DEVELOPMENT

weight, bitrochanteric diameter, calf and chest girths, and waist height comparable with the respective measurements of children in the other places. Analyses of the data by age groups has shown that children in the two southern cities reached a given stature and hip girth, and other measurements, at an earlier age than did those in the northern cities. Hence, it may be inferred that the children in San Antonio and Houston grew at a faster rate between 7 and 14 years than did those of Lubbock and Denton.

Evaluation of Growth

Further comparison of places was made by plotting the mean height and weight for each age group of boys and of girls on Wetzel grids (4, 18). For this examination of the data, two places were plotted on the same chart—San Antonio and Lubbock as one pair, Houston and Denton as another, and San Antonio and Houston, a third.

The position of the successive age groups does not represent the *growth* of any one child or of any group of children, because each child was measured only one time and therefore had a share in representing only one age group. However, considering each age group as a separate unit, its position on the grid does indicate the average body shape and size of that age group of boys or of girls, and shows how size at the given age compares with the speed of development of designated percentiles of the general population.

In all four places both boys and girls were of medium body type, with the 7- and usually the 8-year-old children in the M channel, but older age groups were slightly more slender, placing in the B₁ channel. The size for age of both boys and girls in Lubbock and Denton was closer to the schedule of "standard" development (67 percentile) than were the age groups in San Antonio and Houston. The 7- to 12-year-old groups in Houston were nearer the 15 percentile line than to the 67. Age groups for girls were usually more than 12 developmental levels apart, year to year, from 9 to 12 years, and likewise for boys, from 10 to 12, and again at 13 to 14 years.

These diagrams showed clearly the likeness of body shape and size of children in San Antonio and Houston, and the greater height and weight for children of the same age for these places than for those in Lubbock and Denton. However, it can not be said that the position of the children in San Antonio and Houston is either more or less advantageous than of those in Lubbock and Denton.

Although the idea has been held long and widely that rapid growth and large size within the range considered normal is desirable, it has been seriously questioned by Meredith (6, 7). He has stated that

When the trend toward greater stature and body weight during recent decades is rigorously evaluated against appropriate criteria (that is, criteria representing health, efficiency and longevity), it is possible that it may be found desirable, undesirable, or unrelated to the criteria (6, p. 918).

Rate of children's growth was a point considered in a recent scientific conference (15) where the discussions centered around man, and in particular

around the child, in the matter of supplying the need for protein. The question was raised as to whether a maximal rate of weight gain or tissue formation is necessarily in the long run the best for the organism. The editors of the conference proceedings remark that "from the absence of discussion it would seem that there is not much evidence, one way or the other, on this important point" (p. 69).

Possible Factors Influencing Place Differences in Mean Measurements

Difference in food consumption as the first possible explanation of size differences in the four places is a logical thought. However, this study did not include any information on diet or eating habits. But results from a previous study on the diet of Texas school children (20) in three regions of the state (two of them 75 to 100 miles from Houston) showed no outstanding differences in the kinds of food eaten, nor in the average amount as estimated by the pupils who kept their own daily records. The one difference meriting mention was the greater number of eggs eaten by the children in the two counties where there were relatively more farmers. A survey of the food supply of 400 rural families (21), 15 years later, in three regions (one of which included Lubbock and Lamb Counties and the other two were about 100 and 200 miles from Dallas and Denton, and 200 to 300 miles from Houston) showed a somewhat better provision for nutrients in the Lubbock region where more families were on a higher socioeconomic level than in the other regions. But, assuming that Lubbock families had an excellent food supply also at the time of this body measurement study, the Lubbock children might be expected to have been larger, not smaller, than those in the two southern cities since the four places appeared comparable socioeconomically. Even though it be granted that the San Antonio and Houston children ate more food, or utilized their food better, the question still remains as to *why* they ate more.

Racial stock has long been considered to be a determining factor in body size. With this idea also, Meredith (6) takes issue. He compiled data from a number of studies made with 9- and 14-year-old boys in nine places (as far apart as New York City and Washington, D.C., from Los Angeles and San Francisco, and as Milwaukee from Laredo, Texas) which showed average stature at 9 years to differ less than three-fourths inch among race and nationality groups which included Japanese, Mexican, Pueblo Indian, "pure American" (north European ancestry), Bohemian, Italian, Chinese, Negro, German and Jewish. Meredith observes that the compiled data

are decisive with regard to the untenability of the position that stature varies with race and is fairly rigorously nonadaptive. . . . [If stature is genetically determined] . . . it is necessary to credit "genetic determiners" with having produced like results in boys of Negroid, Mongoloid, and white stocks (p. 923).³

³ Since the completion of this report of the Texas study, William Walter Greulich's article, "Growth of Children of the Same Race under Different Environmental Conditions," appeared in *Science* (1958, 127, 515-516). Greulich found in 1957 that at every age from 6 to 19 years American-born Japanese boys "exceeded in average stature the

CHILD DEVELOPMENT

This study contains no information as to ancestry except for birthplace of both parents of each child. But whatever genetic differences there may have been were without recognizable effect when the data were classified by hip girth classes within stature class; for then place differences disappeared.

Socioeconomic status is well known to be correlated with growth and size as was also found in this study. Meredith (6), in his digest of studies on height and weights of 9- and 14-year-old white boys in the United States, determined that those from the professional and major managerial classes were taller and heavier than those of unskilled and semiskilled classes. He found the magnitude of the differences not to exceed 3 per cent for stature and 6 per cent for weight. In this study stature of boys and of girls in the professional and semiprofessional groups exceeded that of common labor groups by from 1.92 to 4.86 per cent, and averaged 2.76 per cent, in those stature-hip girth classes where the differences were significant. Similarly for weight, these two higher socioeconomic groups were from 6.78 to 15.01 per cent heavier (average, 10.83 per cent) than the common labor group.

Climatic environment has been mentioned in a general way as having an influence on rate of growth and ultimate size. Wetzel *et al.* (19) have observed that "it happens that any one of a number of nonspecific agents, even mere environmental change, can set recovery from growth failure in motion" (p. 652). As early as 1910, Boas (1) presented evidence of environmental influence on growth in that children born in the United States of immigrant parents were taller than children born of the same parents in their native land. Woodbury (22), from data on stature and weight obtained during the Children's Year Campaign of 1918 on white subjects under 6 years, found that those in California were slightly larger than those in eastern and northern states. He considered that his observations were due either to climate or factors related to immigration.

Lloyd-Jones (5), however, has reported definite evidence of influence of environment on growth. This author who measured over 100,000 children in Los Angeles public schools found that they were taller and heavier than public school children in corresponding age groups (6 to 14 years inclusive) in three other California groups and one composite group from several sections of the United States. Los Angeles children were measured in 1936 and 1937; those in Oakland in 1893, those in the "gifted children" group and in the United States composite group in 1923, and those in San Francisco in 1923 and 1924. More impressive, therefore, is the finding that of these Los Angeles children those who were born in California and had lived there continuously (73,258 of them) were taller and heavier, in every year of age group, than were those born elsewhere (40,038 of them) but

boys of Japan by an amount greater than the increase which had taken place in the average stature of the boys of Japan" between 1900 and 1952, and similarly for American-born Japanese girls from 6 to 14 years. Greulich interprets the less advanced skeletal status found for Japanese children in Japan, compared with California Japanese and Caucasian children of Cleveland, Ohio, as more likely due to a less adequate diet and to other environmental conditions than to some characteristic racial difference.

had lived in California 1 to 5 years, 6 to 10 years, and 11 years or more. The longer that children born elsewhere had lived in California, the greater was their average stature and weight. Lloyd-Jones concluded that the superiority in height "revealed among these Los Angeles children must in part be put down to environmental influences in that area" (p. 21).

Consideration of Climatic Factors in the Four Places

Conversation with local animal husbandry men provided the information that physical comfort, as affected by climatic factors, is definitely favorable to the growth and normal performance of animals. For example, rats, if the temperature is uniform, can adjust to a low level, even 41.0°F (5.0°C) after a time. In summer there is poor growth, reproduction, and livability. Chickens and turkeys make their maximum growth at a range of 60° to 70°F (15.56° – 21.11°C). At 85°F (28.33°C) and above, egg laying is cut down and egg shells are softer. Cattle are uncomfortable at temperatures above 90°F (32.22°C), and at 32°F (0.00°C) if the relative humidity is high. Sheep suffer drastic effects from hot weather; weight is barely maintained in summer, unless relative humidity is low; lambs weaned the first of May or later have poor appetites. Swine provided with a cement basin wallow during a 90-day period with average temperature of 85°F (29.44°C) ate more of the ration and gained as much weight in 76 days as their pair mates gained in 90 days on the same ration, but without a wallow.

This information made it seem reasonable to attempt to find any distinguishing characteristics of climate in the four places which might possibly be related to physical comfort of humans, and hence influence growth.

The two places toward the north of the state, Lubbock and Denton, are at latitude 33°N , while San Antonio and Houston are some 225 to 300 miles to the south at 28°N and 29°N latitude respectively. The two northern places are about 300 miles apart; the two southern, 200 miles. Elevations are: San Antonio, 701 ft.; Houston, 55 ft.; Denton, 620 ft.; Dallas, 512 ft.; Lubbock, 3241 ft.

The four places are located in four of the 12 climatic divisions of the state which Russell (10) has mapped using frequencies of climatic years over the period 1914–1931 inclusive to locate boundaries. Houston is near the southwestern edge of the easternmost division which has purely mesothermal temperatures and humid years. Here there is high probability of escaping extreme drought during years when dry climates are most expanded in area. San Antonio lies at the eastern border of the division characterized by mesothermal temperatures and nuclear steppe climate and is separated from the easternmost division by two transitional zones which differ in drought intensity. Dallas lies northward in the first region of transition from humid east Texas to the nuclear steppe to the west. Lubbock is in the steppes of west-central Texas where climate is dry but not arid, and marked temperature changes are common as between day and night,

CHILD DEVELOPMENT

winter and summer, one day and the next. Lubbock and Dallas have in general more continental climate than have San Antonio and Houston.

For details pertaining to local climatic elements, weather records for each place over the span of years 1923-1937 were examined. These years cover the growth periods of the children who were measured.

That the years of children's growth constitute a sufficient period of time to reveal climatic characteristics was determined by analysis of variance to compare the mean annual temperature and total precipitation over the span of 15 years (1923-1937) with the total span of 43 years (1914-1956). Data for these analyses were taken from the United States Weather Bureau's 1956 issue of Local Climatological Data with Comparative Data (16) for San Antonio, Houston, Lubbock, and Dallas. The mean of the mean annual temperatures for the 15-year and the 43-year spans respectively were for San Antonio 69.6 and 69.4° F; Houston, 69.5 and 69.8; Lubbock, 60.4 and 60.0; Dallas, 65.7 and 65.9. Similarly, during the short and the long spans of years the total precipitation was: San Antonio, 27.71 and 26.75 inches; Houston, 42.18 and 44.21; Lubbock, 18.59 and 18.39; Dallas, 34.80 and 33.74. Analysis of variance shows that the 15-year span does provide a representative sample of the climatic conditions in the four places.

The weather data used to compare the four places for the 15 years (1923-1937) were taken from the records of the agricultural experiment

TABLE II

SIGNIFICANT DIFFERENCES (°F) BETWEEN MONTHLY MEAN MAXIMUM AND MINIMUM TEMPERATURES IN ALL POSSIBLE COMPARISONS AMONG THE FOUR PLACES IN WHICH CHILDREN WERE MEASURED

Year	Lubbock more than			Denton more than		San Antonio
	San Antonio	Houston	Denton	San Antonio	Houston	more than Houston
1923	5.03**	6.66**	3.76**			
1924	7.69**	8.22**	4.77**	2.92*	3.45**	
1925	4.53	7.70**			4.68**	3.17*
1926	3.87**	4.30**	2.62*			
1927	9.48**	6.94**	5.67**	3.81**		-2.54*
1928	7.10**	6.30**	3.65**	3.45*		
1929	7.72**	7.98**	5.03**	2.69*	2.95*	
1930	7.96**	8.21**	4.11**	3.85*	4.10**	
1931	6.81**	5.85**	3.83**	2.98**		
1932	7.10**	6.85**	3.56**	3.54*	3.29*	
1933	9.82**	10.95**	6.87**	2.95*	4.08**	
1934	9.65**	9.69**	6.17**	3.48**	3.52**	
1935	8.90**	7.32**	6.00**	2.90*		
1936	8.14**	6.98**	3.67**	4.47**	3.31*	
1937	6.73**	5.76**	4.61**			

NOTE.—Weather data for Houston from Angleton which is the agricultural experiment station nearest to Houston.

* Significant at 5% level.

** Significant at 1% level.

substation at each place. However, for Houston, the records for Angleton, the nearest substation (about 45 miles away) were used.

The differences between mean monthly maximum and mean monthly minimum temperatures in the four places were compared, year by year, by analysis of variance. The results are shown in Table 11 for the six possible pairs of places. The spread between monthly maximum and minimum mean temperatures at Lubbock was significantly greater, beyond the .01 level, than such differences at San Antonio and at Houston in every year, and between Lubbock and Denton, significant at the .01 level in 12 years, but significant at the .05 level in 2 years. Maximum-minimum differences between Lubbock versus San Antonio and Lubbock versus Houston were strikingly similar, ranging (in 1926) from 3.87°F (San Antonio) and 4.30°F (Houston) to 9.82°F (San Antonio) and 10.95°F (Houston) in 1933. Extreme difference in monthly mean temperature in Lubbock exceeded that in Denton by values of 2.62°F (1926) to 6.87°F in 1933. Extreme differences of temperature in Denton were significantly greater than in San Antonio for 11 years, (seven of them at .05 level); and greater than in Houston in eight years. But San Antonio's excesses over Houston were significant in only two years, with opposite sign for the small values of -2.54° and $+3.17^{\circ}\text{F}$. Clearly, the two northern places, Lubbock and Denton, were subject to far greater extremes of temperature than were the southern places, San Antonio and Houston.

The greater prevalence of lower temperature in Lubbock and Dallas (Denton) is shown also by their higher number of heating degree days. Degree days are found by subtracting from 65°F each daily mean temperature which is below 65° . The sum of such differences for the days within any month expresses the degree days for that month. The sum of monthly values within a year gives the annual degree days. The degree days over the 30-year period 1921-1950 as given in the U.S. Weather Bureau summary (17) of monthly normal temperature, precipitation, and degree days are: Lubbock 3587, Dallas 2272, San Antonio 1579, Houston 1388. However, Dallas, like San Antonio and Houston, had no degree days during the five months of May through September, while the no degree days for Lubbock included only June, July, and August.

But temperature does not act alone in its effect on physical comfort. Thornthwaite (14) has stated that "the sum or climatic elements that have been under observation does not equal climate" (p. 55). Two of the climatic factors that have been considered in conjunction with temperature are wind velocity and relative humidity. Siple (11, 12) coined the expression "wind chill" to describe the cooling effect of air movement and designated temperature. At a given difference between a comfortable temperature of uncovered skin and air temperature, the greater the wind velocity, the greater the body discomfort unless suitable clothing is worn. Lacking observations of wind velocity and temperature known to have been made simultaneously—essentials according to Court (3)—and at the same hour of day at the four stations, no calculations of wind chill factors have been made.

CHILD DEVELOPMENT

It is interesting to note, however, that with lower temperatures at Lubbock and Denton and also higher wind run totals and wind chill values would likely be greater in these places than for San Antonio and Houston. What influence mental attitude might have in adjusting (3, 11) to wind chill values in the different places is not known. Thom (13) has observed that the same amount of discomfort occurs under widely differing dry-bulb temperature conditions. He cites the illustration that the same maximum discomfort index of 81°F was produced at Hatteras, North Carolina, by a dry-bulb temperature of 84°F and wet-bulb of 80°F (relative humidity 84%), as at Sheridan, Wyoming, where the dry-bulb temperature was 101°F , and wet-bulb, 64°F (relative humidity 12%). It was not possible to calculate discomfort indices for the four places in this study, because the necessary data were not available from either the station or U.S. Weather Bureau records for the years 1923-1937.

In this study a comparison by graphs was made of the course of the concurrent monthly mean temperature and mean relative humidity through

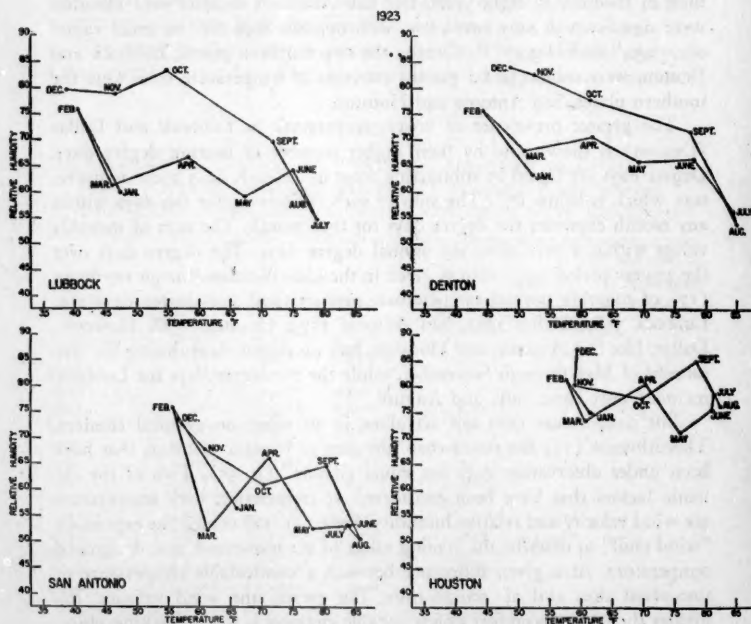


FIGURE 11—Climogram showing monthly means of concurrent relative humidity and temperature for the year 1923, in the four places in Texas where children were measured.

the year. A climogram (2) was drawn for each of the 15 years. Representative climograms are shown in Figures 11, 12, 13, and 14. The years to use for illustration were chosen on the basis of the magnitude of the difference between monthly mean maximum and minimum temperatures. Houston, Denton, and Lubbock had their smallest such difference in 1923, while San Antonio was close to its smallest in that year. In 1925 San Antonio and Denton had their greatest extreme difference of monthly mean maximum and minimum temperatures, while Houston and Lubbock were near their greatest value. The years 1932 and 1936 represent midway conditions of monthly mean temperature range for the respective places.

In each year, as seen in the climograms, there seemed to be a general pattern applicable to all four places; but within years there were consistent place differences.

Considering each year as a unit, San Antonio had the lowest relative humidity, Houston the highest. Excess of relative humidity in Denton over that in Lubbock applied more, in general, to the months of March

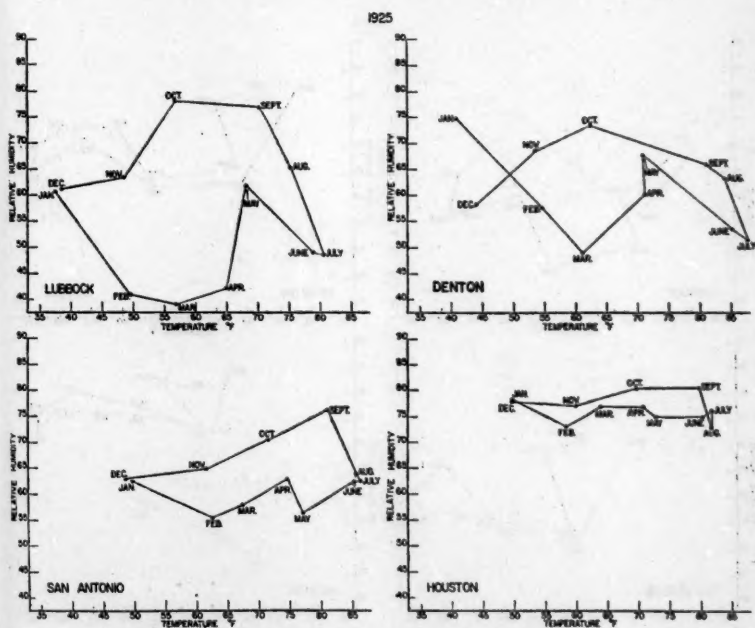


FIGURE 12—Climogram showing monthly means of concurrent relative humidity and temperature for the year 1925, in the four places in Texas where children were measured.

CHILD DEVELOPMENT

through July than September through December. The year 1932 was exceptional in Denton in that the relative humidity was lower (instead of higher) from September through November than from March through June.

The lower monthly mean temperatures in Lubbock and Denton than in San Antonio and Houston would appear, by the climograms, to be due largely to their lower temperatures during the months of January, February, November, and December; while at the high end of the temperature range there was less difference between the northern places (especially Denton) and the southern places during the months of May through August.

In every year the climograms for Lubbock and Denton enclosed a larger area than did those for San Antonio and Houston in the same year. On the whole, the course of concurrent relative humidity and temperature through the year was more widespread in Lubbock and Denton than in San Antonio and Houston.

The cumulative totals of each of several climatic factors over the 15-year period are shown in Table 12. Mean temperatures at both San Antonio and

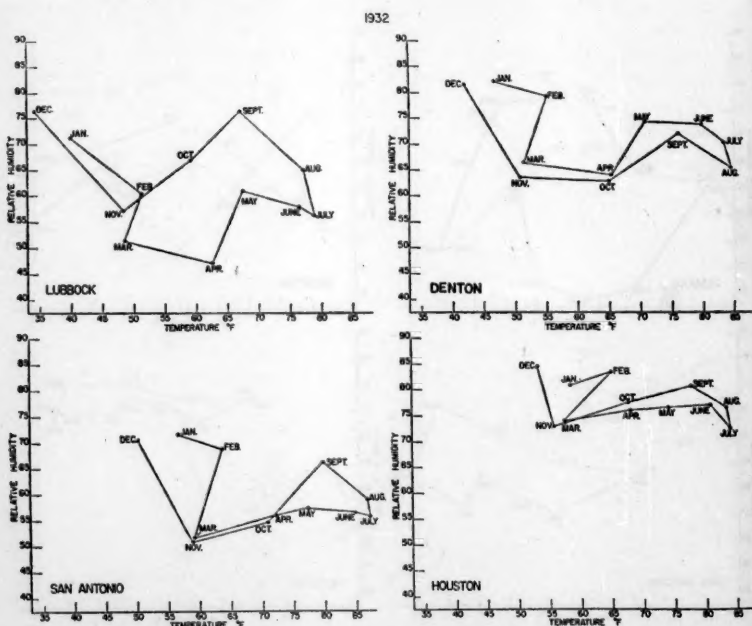


FIGURE 13—Climogram showing monthly means of concurrent relative humidity and temperature for the year 1932, in the four places in Texas where children were measured.

JESSIE WHITACRE and ETHEL T. GRIMES

Houston were significantly greater than at Lubbock and at Denton, while Denton temperature was significantly greater than at Lubbock. Nearly twice as many killing frosts were recorded at Lubbock as at Denton and approximately five times as many as at San Antonio and Houston.

Lubbock had the least precipitation, San Antonio and Denton similar amounts which were significantly greater than at Lubbock, but significantly less than at Houston.

Denton with the highest wind run and Lubbock in second place both differed significantly from Houston and San Antonio. Wind run at Houston was significantly greater than at San Antonio.

Although character of days was not determined objectively, it is interesting that cloudy days varied little among the four places; but in clear days Denton was significantly ahead of Lubbock and both those places ahead of San Antonio and Houston, and San Antonio, in turn, ahead of Houston. Excess in number of wind and dust storms, hot winds, and northers at Lubbock and Denton was far greater than for the other weather factors.

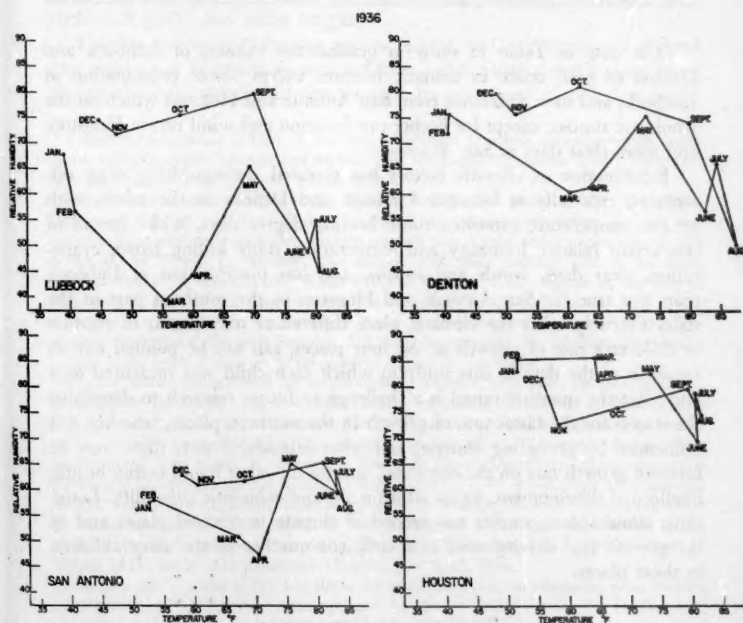


FIGURE 14—Climogram showing monthly means of concurrent relative humidity and temperature for the year 1936, in the four places in Texas where children were measured.

CHILD DEVELOPMENT

TABLE 12

SOME CHARACTERISTICS OF CLIMATE AT THE FOUR PLACES: TOTALS
OVER THE 15 YEARS, 1923-1937, INCLUSIVE

<i>Characteristic</i>	<i>San Antonio</i>	<i>Houston</i>	<i>Lubbock</i>	<i>Denton</i>
Temperature (°F) mean monthly . . .	1061.6	1032.5	905.5	972.3
Total precipitation (inches)	442.7	675.2	278.9	457.3
Evaporation (inches)	no data	663.3	961.7	881.8
Clear days/month (number)	2247	1622	2800	3023
Cloudy days/month (number)	1284	1318	1208	1200
Fog (number)	230	667	113	155
Frost, killing (number)	196	222	1244	747
Wind run (miles)	395426	603682	840631	894088
Wind storms (number)	54	35	885	253
Dust storms (number)	30	10	691	120
Hot winds (number)	16	0	234	91
Northers (number)	254	355	659	511

The data in Table 12 show in general the likeness of Lubbock and Denton to each other in climatic features except lower precipitation at Lubbock, and their difference from San Antonio and Houston which on the whole are similar, except for higher precipitation and wind run at Houston, and more clear days at San Antonio.

Examination of climatic factors has revealed distinguishing place differences, especially as between Lubbock and Denton to the north, with greater temperature extremes, more heating degree days, wider spread of concurrent relative humidity and temperature, more killing frosts, evaporation, clear days, winds and storms, and less precipitation at Lubbock than was true for San Antonio and Houston in the southern part of the state. Precisely what the climatic place differences may mean, in relation to children's rate of growth in the four places, can not be pointed out on the basis of the data in this study in which each child was measured only once. But the question raised is a challenge to future research to determine the reason for the faster rate of growth in the southern places, whether it is influenced by prevailing climate, and what relation, if any, there may be between growth rate on the one hand, and on the other hand, bodily health, intellectual development, social adjustment, and economic capability. Long-time, simultaneous studies are needed of climate in selected places and of the growth and development of a sufficient number of the same children in those places.

SUMMARY AND CONCLUSIONS

Body size of 3000 boys and 3535 girls of 7 to 14 years of age has been studied in relation to place of residence and occupational classification of

JESSIE WHITACRE and ETHEL T. GRIMES

their parents.⁴ Means of the eleven measurements made at one time on each child showed for all ages combined, and for each year of age separately, that the children in San Antonio and Houston were larger in every measurement than were those in Lubbock and Denton. The places to the north, Lubbock and Denton, were 225 to 300 miles distant from the southern ones, San Antonio and Houston. The northern places were some 300 miles apart, the southern 200 miles. Lubbock had greatest elevation; Denton lowest population.

Means of stature, hip girth, bitrochanteric diameter, and weight, all ages together, were significantly larger (at .01 level) for both boys and girls in the professional, semiprofessional and business groups than those in the skilled and common labor groups whose means resembled each other. Farmer boys all ages combined, resembled labor groups, but farmer girls, the semi-professional. Single year-of-age groups showed trends among occupational groups similar to those for all ages combined.

When age was disregarded and the children were classified by hip girth intervals within stature intervals, place differences disappeared in the means of boys and girls alike for weight, bitrochanteric diameter, chest girth, calf girth, and waist height.

Heights and weights for age groups plotted on the Wetzel grid showed that children in all places were of medium build, those above 8 years being slightly the more slender. In size for age, the age groups placed close to the

⁴ Sincere appreciation is held of the fine cooperation of school superintendents, principals, teachers, pupils' parents and especially of the pupils themselves which made possible the study of children's body measurements.

Indebtedness is acknowledged to the specially trained field workers, Mary Bess Egan, Laverne McWhirter, and Alyce Strange, and their co-workers in measuring, for their sustained interest and meticulous care in their exacting task; to the anthropometrists, Charles E. Snow and Everett L. Marshall who assisted with their training; and to the many women who, in service with the Works Progress Administration, assisted at the time of measuring as recorders of measurements and in caring for the children.

Gratitude is expressed to the Bureau of Home Economics (now Institute of Home Economics, Agricultural Research Service), U.S.D.A., for granting the privilege of retaining a copy of the original data on Texas children for use in this study.

Cordial thanks are extended for information on weather items from the men in charge of certain of the United States Weather Bureau stations in Texas—O. E. Edrington at San Antonio, E. A. Farrell at Houston, Harold Frost at Lubbock, and M. E. Harrison at Dallas. Greatly appreciated help, through personal counsel and literature references, has been received from Arnold H. Glaser and Walter K. Henry, Department of Oceanography, Texas A & M College, College Station, Texas; Richard D. W. Blood, State Climatologist, Austin, Texas; Earl C. Thom, Meteorologist, U. S. Weather Bureau, Washington, D.C.; L. T. Pierce, State Climatologist, U. S. Weather Bureau, Columbus, Ohio; Robert Basile, Department of Geography, Ohio State University, Columbus, Ohio; and Norman C. Wetzel M.D., Babies and Children's Hospital, Cleveland, Ohio.

Credit is gladly given to Ivy Lee Beene for excellent service in tabulating data, making calculations, and typing the manuscript, and for similar help to Helen M. Darling who also drew graphs.

Most grateful acknowledgment is made to C. B. Godbey and Robert R. Shrode, Department of Genetics, Texas A & M College, College Station, Texas, for their guidance and generous assistance with the statistical treatment of the data.

CHILD DEVELOPMENT

schedule of "standard" development (67 percentile) except that the 7- to 12-year-old groups in Houston were nearer the 15 percentile than the 67. Age groups for girls were usually more than 12 developmental levels apart, year to year, from 9 to 12 years, and likewise for boys from 10 to 12, and again 13 to 14 years. The grids showed clearly the greater height and weight for age of the San Antonio and Houston children than of those in Lubbock and Denton. It cannot be said that the physical development of the children in any one place is more or less advantageous than in the others, as optimal rate of growth is not now known.

As possible factors to explain place differences in mean measurements, influence of children's age, possible difference in food consumption, possible difference in genetic background and in socioeconomic levels seemed to be ruled out.

Characteristic place patterns of climate were found respecting mean temperature, relative humidity and temperature combinations, precipitation, evaporation, clear and cloudy days, winds and storms. In general, these patterns distinguished the two northern places from the two southern ones. The correspondence between climatic patterns and place difference in rate of growth may be coincidental. But these findings present a challenge for future research to determine the reason for the place difference in rate of growth, whether or not rate is influenced by climate, and what relation growth rate may have to health, intellectual development, social adjustment, and economic capacity.

REFERENCES

1. BOAS, F. *Changes in bodily form of descendents of immigrants.* (Senate Document No. 208) Washington, D.C.: U. S. Government Printing Office, 1910.
2. CONRAD, V. *Methods in climatology.* Cambridge, Mass.: Harvard Univer. Press, 1944.
3. COURT, A. Wind chill. *Bull. Amer. meteorol. Soc.*, 1948, 29, 489-493.
4. *Growth measurement. A synopsis of the Wetzel grid technique for evaluating the quality of physical growth and development in pre-school and in school children.* Abbott Laboratories, 1946.
5. LLOYD-JONES, O. California tall children. *Amer. J. Dis. Child.*, 1940, 60, 11-21.
6. MEREDITH, H. V. Stature and weight of children of the United States with reference to the influence of racial, regional, socioeconomic and secular factors. *Amer. J. Dis. Child.*, 1941, 62, 909-932.
7. MEREDITH, H. V. Stature and weight of private school children in two successive decades. *Amer. J. phys. Anthropol.*, 1941, 28, 1-40.
8. O'BRIEN, RUTH, & GIRSCHICK, M. A. *Children body measurement for sizing garments and patterns.* U. S. Dept. Agric., Misc. Publ., No. 365, 1939.

JESSIE WHITACRE and ETHEL T. GRIMES

9. O'BRIEN, RUTH, GIRSHICK, M. A., & HUNT, ELEANOR P. *Body measurements of American boys and girls for garment and pattern construction*. U. S. Dept. Agric., Misc. Publ. No. 366, 1941.
10. RUSSELL, R. J. *Climates of Texas*. *Ann. Ass. Amer. Geographers*, 1945, 35, 37-52.
11. SIPLE, P. A. *Climate and clothing*. Chapter 12 in *Physiology of heat regulation*. Philadelphia: Saunders, 1949.
12. SIPLE, P. A., & PASSEL, C. F. *Measurements of dry atmospheric cooling in subfreezing temperatures*. *Proc. Amer. Philosophical Soc.*, 1945, 89, 177-199.
13. THOM, E. C. *A new concept for cooling degree-days*. *Air Conditioning, Heating & Ventilating*, 1957, 54, 73-80.
14. THORNTHWAITTE, C. W. *An approach toward a rational classification of climate*. *Geographical Rev.*, 1948, 38, 55-94.
15. WATERLOW, J. C., & STEPHEN, JOAN M. L. *Human protein requirements and their fulfillment in practice*. New York: Josiah Macy Jr. Found., 1957.
16. Weather Bureau, U. S. Dept. of Commerce. *Local climatological data with comparative data. For each of four stations—San Antonio, Houston, Lubbock, Dallas*. 1956.
17. Weather Bureau, U. S. Dept. of Commerce. *Monthly normal temperature, precipitation and degree days*. Tech. paper No. 31, 1956.
18. WETZEL, N. C. *Instruction manual in the use of the grid for evaluating physical fitness*. Cleveland: NEA Service, 1941.
19. WETZEL, N. C., FARGO, W. C., SMITH, ISABEL H., & HELIKSON, JOSEPHINE. *Growth failure in school children as associated with vitamin B₁₂ deficiency—response to oral therapy*. *Science*, 1949, 110, 651-653.
20. WHITACRE, JESSIE. *The diet of Texas school children*. *Texas Agric. Exp. Sta. Bull.*, 1934, No. 489.
21. WHITACRE, JESSIE. *The food supply of Texas rural families*. *Texas Agric. Exp. Sta. Bull.*, 1943, No. 642.
22. WOODBURY, R. M. *Stature and age of children under 6 years of age*. Children's Bureau, U. S. Dept. of Labor, 1921, No. 87.

THE RELATIONSHIP OF EARLY INFANT REGULATION AND LATER BEHAVIOR IN PLAY INTERVIEWS

MARY ELLEN DURRETT*¹

San Jose State College

Experiences of early infancy and childhood are alleged to be crucial in the formation of personality and in the etiology of behavior deviations. Much attention has been paid to the infant's first social contacts, those with his parents, most of them involving the handling of his feeding, sleep, and elimination needs. Sewell and Mussen (6) suggest that these relationships are probably the most important aspects of the infant's early environment, and that early treatment and care must somehow be influential in determining the course of the child's later personality development, although the relationship between specific techniques of child guidance and particular personality traits may not be as direct or as great as many psychoanalytically-oriented writers and others believe. The available data do not provide unequivocal evidence of the exact nature in which early life conditions are related to later development.

The present study was initiated to investigate the relationship between fantasy aggression of 4- and 5-year-old children and early child-rearing experiences. It was hypothesized that there is a positive relationship between such aggression and strictness of regulation during infancy so that, if a mother uses very strict child-rearing practices, the child will display a high level of aggression.

PROCEDURE

Subjects

Sixty children ranging in age from 4 to 6 years, and their mothers, from white, native-born, urban, intact families living in Tallahassee where the father followed a profession, served as subjects for this study. Of the 60 families represented, 50 of the fathers were members of the faculty at Florida State University and 10 followed a profession such as law or medicine. According to the Index of Social Status—Short Form, as devised by McGuire and White (2), which is based on occupation, source of income and education, 58 of the families were in the upper middle class, and two were in the lower upper class.

The children were the biological offspring of both parents with whom they lived, i.e., families with either adopted or stepchildren were not included in the study. There were 15 male and 15 female 4-year-olds, and

* San Jose State College, San Jose 14, California.

¹ A condensation of a dissertation presented in partial fulfillment of the requirements for the degree of Doctor of Philosophy, Florida State University, June, 1957. The author wishes to express her gratitude to her graduate committee: Drs. Ruth Connor, Ralph Witherspoon, Ruth J. Dales, James Walters, and Charles M. Grigg.

CHILD DEVELOPMENT

15 male and 15 female 5-year-olds. The mean age was 4 years, 10 months, the range being 4 years to 5 years, 10 months, with a standard deviation of 6.6 months. None of the children had physical handicaps nor severe health problems; none showed signs of mental deficiency; none had had previous experience with the doll play interview.

Measurement of Fantasy Aggression

Data were obtained in controlled, standardized doll play interviews, following the procedure described by P. S. Sears (4). Each child was given two play interviews of 20 minutes each, never on consecutive days, but always within a 5-day period. The categories used were the same or a modification of those used in previous doll play studies (1, 4). These aspects of behavior were recorded: (A) indicated physical aggression—an action that had the intent to irritate, hurt, injure, punish, frustrate, or destroy a doll or equipment. (a) indicated verbal aggression—a verbal expression that had the intent to irritate, hurt, injure, punish, frustrate, or destroy a doll or equipment; scolding, threats, and uncooperativeness were included as were instances in which discomfort was attributed to a doll ("the boy is sick, sad, lost, etc."); (-) indicated nonaggression—any behavior which was not included in the above categories. The child's behavior was recorded in terms of the above symbols every 15 seconds. If verbal aggression or physical aggression or both occurred within any 15-second interval, it was recorded. The measure used in the present study was the "percentage of intervals" in which aggression occurred as contrasted with intervals in which aggression did not occur.

After approximately 25 hours of preliminary practice with the recording procedure, reliability was computed on the basis of 15 20-minute sessions of doll play with two observers, using children from the nursery school and kindergarten who were not serving as subjects in the investigation. One observer sat behind a one-way vision mirror and recorded the children's behavior independently of the experimenter who conducted and recorded the experimental sessions. Reliability was computed by means of percentage of agreement between the two observers, using the formula: $2(\text{number of agreements}) / \text{total number of observations}$. In order to constitute an agreement, not only the same symbol or symbols had to be recorded by the two observers, but also within the same 15-second interval. The percentage of agreement between the observers based on these 15 20-minute sessions for total aggression, physical aggression, verbal aggression, and nonaggression was 88, 83, 80, and 97, respectively.

Measurement of Early Regulation

After the completion of the doll play sessions, the children's mothers were interviewed in their homes and asked to rate themselves on five scales, modeled after those used by Sears and his collaborators (5): one rating on the child's feeding schedule—the degree of strictness, from an exact clock schedule to complete self-regulation; two scales concerning toilet training—the age when it was begun, and the way toilet accidents were

MARY ELLEN DURRETT

handled; and two ratings on discipline and guidance—the amount of obedience expected of the child, and the way of handling aggression toward parents. Each 5-point scale included responses ranging from very permissive to very strict, although the responses were not presented to the mother in this order. Later, the mother's ratings were converted into a scaled score, using a key which included the same responses found on the interview schedule, but arranged from very permissive to very strict.

Ten weeks after the interviews were completed, 15 of the mothers rated themselves a second time. The mean percentage of agreement between the two sets of responses was 92.

RESULTS

Fantasy Aggression

When the difference in the number of 15-second intervals in which aggression was shown for session 1 and session 2 was tested for significance, a *t* of 11.81 was found, significant at the .001 level. Forty-five of the differences were increases from the first to the second session, the children showing significantly more aggression in the second session. The correlations between session 1 and session 2 for the percentage of intervals in which aggression was shown for boys, girls, and total group, respectively, were .59, .60, and .66, all significant at the .001 level. The scores for the two sessions for each child were therefore combined to obtain an over-all measure of fantasy aggression. The results are presented in Table 1.

TABLE 1
SEX DIFFERENCES IN AGGRESSION

	N	PERCENTAGE OF INTERVALS SHOWING AGGRESSION	
		Physical	Verbal
Boys	30	82.0	49.7
Girls	30	37.6	86.3
		$p < .01$	$p < .01$

It will be noted that boys showed a significantly higher percentage of physical aggression, out of the total number of intervals in which aggression occurred, than did girls. However, girls showed a significantly higher percentage of verbal aggression than did boys. The percentage of intervals in which either verbal or physical aggression was shown was significantly higher for boys than for girls. For both sexes, there was a significant session-to-session increase in the number of intervals in which aggression, physical, verbal, either or both, was shown. There was also a significantly higher

CHILD DEVELOPMENT

percentage of intervals in which aggression occurred among 5-year-olds than among 4-year-olds.

Early Regulation

Intercorrelations of measures of early regulation were low and positive with few exceptions. Only two of the coefficients were significant at the .05 level: between degree of regulation of feeding schedule and timing of toilet training for boys; and between degree of regulation of feeding schedule and way of handling toilet accidents for girls. The data also disclosed a significant sex difference in frequency of spanking by mother and by father, as shown in Tables 2 and 3.

TABLE 2
DIFFERENCES IN FREQUENCY OF SPANKINGS GIVEN BOYS AND GIRLS BY FATHER

Frequency	Boys N=30	Girls N=30
1. About once a week, more than once a month, less than once a week ..	9	4
2. More than twice a year, not more than once a month	7	9
3. Rarely, once or twice a year	11	5
4. Once, twice, or three times in lifetime, or never	3	12
$\chi^2 = 9.82$	$p < .05$	

TABLE 3
DIFFERENCES IN FREQUENCY OF SPANKINGS GIVEN BOYS AND GIRLS BY MOTHER

Frequency	Boys N=30	Girls N=30
1. Several times a week or about once a week	9	2
2. More than once a month, less than once a week	10	7
3. More than twice a year, not more than once a month or rarely (once or twice a year), more frequently when child was younger	6	11
4. Rarely (once or twice a year), not more frequently when child was younger, or once, twice or three times in lifetime or never	5	10
$\chi^2 = 8.12$	$p < .05$	

Fantasy Aggression and Early Regulation

None of the correlation coefficients between aggression and early regulation measures was significant. Only two of the coefficients were in the predicted direction; the highest was +.29, between aggression and way of handling aggression toward parents in boys.

DISCUSSION

The doll play aggression data revealed more physical aggression in boys and more verbal aggression in girls. Although the girls at these ages appeared to be nearly equal in size and strength to the boys, they did not exhibit as much fighting or violence in their doll play. They were more likely to use words to express their aggression; the aggression they did show tended to be in the form of disparagement, scolding, or other nonphysical punishments. It is therefore interesting to note that the boys were spanked more frequently by both the mother and the father than were the girls.

Boys exhibited reliably more total aggression than did girls, thus confirming previous studies (1, 4). Since earlier investigations used populations of somewhat lower socioeconomic status than the present sample, it would appear that this is a reasonably stable finding.

For both sexes, there was a significant session-to-session increase in the number of intervals in which aggression occurred. This finding is also consistent with that of other studies (1, 4) and has been interpreted to mean that the deliberate permissiveness of the experimental procedure progressively acts to reduce inhibitions of aggression that the child has heretofore acquired.

There are several possible sources of explanation for the failure to find a relation between early regulation measures and aggression. The definition and measurement of variables are subject to some of the criticisms presented by Martin (3). For example, the antecedent variables may not have adequately accounted for the meaning of early regulation practices for a particular child. Other variables, such as family constellation, ordinal position, influence of siblings, and mothers' feelings about child training practices, were not measured and not controlled. It may be, also, that the mother is not the best informant about her practices and those of other family members. Most of the mothers in this sample had had a college education and were acquainted with popularized versions of child psychology. Possibly, they were aware of what was considered appropriate behavior and rated their practices accordingly. Furthermore, upper middle class mothers may not show enough range of behavior to produce marked differences in aggression in children. The mothers may not have been able to recall accurately what actually happened in the child training situation. Therefore, the actual practices of the mothers may have differed from what was believed or reported to have happened.

The low intercorrelations of the early regulation measures do indicate that the mothers studied were not consistent in the degree of permissiveness or severity used from one child-rearing situation to another and are similar to those reported by Sewell, Mussen, and Harris (7).

SUMMARY

The general aim of this study was to investigate the relationship between fantasy aggression in young children and the degree of regulation in

CHILD DEVELOPMENT

their infant feeding schedules, the timing of and methods used in their toilet education, and the methods of discipline and guidance used in specific situations.

The subjects were 60 4- and 5-year-old children and their mothers from intact families in Tallahassee in which the fathers followed a profession. The frequency of aggression was measured in two 20-minute sessions of standard doll play interviews with each of the 30 boys and 30 girls. The early regulation data were ratings obtained in interviews with the children's mothers following the completion of the two doll play sessions.

Boys exhibited significantly more physical aggression than did girls, while girls displayed significantly more verbal aggression than did boys. The frequency of spankings given boys by mothers and fathers was significantly greater than that given girls. The findings of earlier investigations were confirmed in that, in both doll play sessions, boys were more aggressive than girls, and for both sexes, there was an increase in the amount of aggression from the first to the second session. No statistically significant relationship was found between aggression and the early regulation measures. Intercorrelations among early regulation measures were, with few exceptions, low and positive.

REFERENCES

1. LEVIN, H., & SEARS, R. R. Identification with parents as a determinant of doll play aggression. *Child Developm.*, 1956, 27, 135-153.
2. MCGUIRE, C., & WHITE, G. D. The measurement of social status. Research paper in human development no. 3 (rev.). Austin: Univer. of Texas, 1955.
3. MARTIN, W. E. Effects of early training on personality. *Marriage fam. Living*, 1957, 19, 39-45.
4. SEARS, PAULINE S. Doll play aggression in normal young children: influence of sex, age, sibling status, father's absence. *Psychol. Monogr.*, 1951, 65, No. 6.
5. SEARS, R. R., MACCOBY, ELEANOR E., & LEVIN, H. *Patterns of child rearing*. Evanston, Ill.: Row, Peterson, 1957.
6. SEWELL, W. H., & MUSSEN, P. H. Infant training and the personality of the child. *Amer. J. Sociol.*, 1952, 58, 150-159.
7. SEWELL, W. H., MUSSEN, P. H., & HARRIS, C. W. Relationships among child training practices. *Amer. sociol. Rev.*, 1955, 20, 137-148.

THE DEVELOPMENT OF THE EXPECTATION OF THE NONINDEPENDENCE OF RANDOM EVENTS IN CHILDREN

GERALD GRATCH*¹

University of Minnesota

Many individuals, untrained in statistics, tend to predict that random events will *never* alternate regularly, e.g., HTHHT will never occur. Further, they tend to expect that the relative frequency of a short sequence of random events will match the probability of the events, e.g., five heads will occur in 10 flips of a coin. Such observations have become salient because the recent emphasis upon statistical reasoning has led to a renewed interest in the problem of measuring "subjective" probabilities or "degrees of belief" with respect to probable events (2). In particular, many investigators have attempted to find order in how people react to random events by comparing their behavior with the statisticians' probability models, e.g., game theorists.

Working within this tradition, Lawlor (10) has impressively documented the anecdotal results mentioned in the previous paragraph. Using adolescents and young adults as Ss, he made a detailed study of their tendency to expect random events to be sequentially dependent. In one portion of his investigation, the Ss estimated the relative likelihood of occurrence of a number of sequences of heads and tails (there were 16 events in each sequence). He found that the Ss guessed that there would tend to be as many runs as possible, with the qualification that regularly alternating sequences, e.g., HTHT . . . , HHTTHHTT . . . , were very unlikely to occur. The sequences that the Ss judged to be the most probable

*Illinois Neuropsychiatric Institute, University of Illinois, Chicago.

¹ This study was adapted from a dissertation submitted to the University of Chicago in partial fulfillment of the requirements for the degree of Doctor of Philosophy. The author wishes to express his appreciation to Professors Helen L. Koch, Lyle V. Jones, and Frederick L. Strodbeck for their advice and help in carrying out this research. Thanks are also due Robert A. Bassham, Prof. Raymond O. Collier, Prof. Henry W. Riecken, and Dr. William Rozeboom for their aid in the preparation of this manuscript. Finally the author wishes to thank Bernard Berkin for making it possible to study the subjects of this study.

CHILD DEVELOPMENT

were characterized by runs seldom longer than two, many runs, and approximately equal numbers of heads and tails. Lawlor's results are consistent with Skinner's (15) analysis of the Zenith experiments on mental telepathy in which a radio audience predicted the outcome of a sequence of five flips of a coin. Therefore, the patterns Lawlor describes would seem to characterize adults' expectations of sequences of binary events generated by a random device such as a coin.

Relatively few attempts have been made to ascertain children's ideas about probable events (1, 13). However, many psychoanalytic writers have commented upon the difficulty that children have in conceiving that events, particularly emotionally significant events, may have an uncertain frequency of occurrence. The present study investigates children's expectations about sequences of random binary events in two ways. First, it tests the simple hypothesis that, with increasing age, children will develop the kinds of sequential expectations that Lawlor found adolescents and young adults to have. Secondly, this study considers the possibility that there is a development with age in the tendency of children to utilize information about the composition of a population of binary events when they are guessing about samples of events being drawn from the population.

METHOD

Subjects. Eighty-four children from an urban public school served as subjects in the study. There were 28 who were in the first grade, 28, in the third grade, and 28, in the fifth and sixth grades. The ages of the children in the three grades were approximately 6, 9, and 11 years. Half the children at each grade level were boys. Only the age comparisons will be considered in this report because no differences were found between the sexes.

Materials. The experimental material consisted of 12 decks of nine cards each. There were two kinds of decks. Six will be referred to as the 5-4 because they were composed of five cards marked with a blue square and four cards marked with a blue diamond (the square and the diamond were equal in area). The other six decks will be referred to as the 7-2 because they were composed of seven cards marked with a blue square and two cards marked with a blue diamond. The cards were of ordinary playing card size. (Hereafter, square and diamond will be referred to by *s* and *d*.)

Procedure. Each child had the following experiences. The child was taken from his classroom and was escorted to a room in the basement of the public school by the experimenter. The child was then seated at a table with the *E* across from him. First, the *E* spread the 5-4 deck face-up in front of the child and then the 7-2 deck. The *E* then told the child how many *ss* and *ds* each deck contained. Next, the *E* showed the child how the decks generated events. He picked up the 5-4 deck and shuffled it well. He took the top card off and showed it to the child. In every case the first card shown was a *d*. The *E* repeated this demonstration four more times

GERALD GRATCH

with the 5-4 deck. The pattern of events thus displayed to each child was *dssds*. The procedure was repeated with the 7-2 deck, and this time the sequence of events was *ssds*. Both decks were then spread face-up in front of the child again, and they remained there throughout the experiment.

The actual guessing task of the child proceeded in the following way. The *E* would take a deck of cards, tell the *S* how many *s*s and *d*s the deck contained, and would ask the child to say whether the top card would be a *s* or a *d*. (The order in which *s* and *d* was said in the instructions was alternated randomly from trial to trial.) After the child guessed, the top card was not shown but the deck was placed in front of the child, another deck was selected, and the procedure was repeated. First, the child guessed the top card in six successive 5-4 decks. Then, he guessed the top card in six successive 7-2 decks. The child recorded his guesses for himself on a sheet of paper. (The *E* did this for the *S*s of age 6.) Then, the child was asked why he guessed as he did, and finally, he was shown the 12 events.

TABLE 1

THE NUMBER OF RUNS IN THE GUESSES OF THE CHILDREN WHO GUESSED THREE SQUARES AND THREE DIAMONDS

Age (years)	NUMBER OF RUNS					
	5-4 DECKS			7-2 DECKS		
	$r \geq 5$	$r < 5$	Total	$r \geq 5$	$r < 5$	Total
6	11	6	17	11	4	15
9	9	5	14	12	4	16
11	5	16	21	2	8	10
Total	25	27	52	25	16	41
$\chi^2 = 8.31, p < .02, 2 df$			$\chi^2 = 26.23, p < .001, 2 df$			

RESULTS

Two kinds of results will be reported. Results based upon the sequential properties of the *S*s' guesses will be presented first, and then the frequency with which the *S*s guessed *s* will be reported.

Table 1 compares the sequences of guesses of the three age groups for those *S*s who chose *s* three times on either the 5-4 or the 7-2 decks.² The important

² There are some problems in the analysis of sequences of events that are introduced by the fact that the probability of the number of the runs and the maximum run length varies with the number of events of each kind in the sequence. Because so many *S*s chose *s* three times, these problems can be by-passed by considering only those *S*s who chose *s* three times. A separate analysis will be performed upon those *S*s who chose *s* a different number of times.

CHILD DEVELOPMENT

comparison is the number of Ss at each age level who guessed many runs or few runs. (A run is a maximal subsequence of like events.) This comparison is made by dividing the Ss into those whose guesses, on a given kind of deck, contained at least five runs and those whose guesses contained less than five runs. (The maximum number of runs in the guesses on the 5-4 decks and on the 7-2 decks is six and the minimum number of runs is two.) The result of this comparison is highly significant. The Ss of ages 6 and 9 tended to alternate their guesses much more often than the Ss of age 11 on both the 5-4 and the 7-2 decks.

TABLE 2

A COMPARISON OF THE OBSERVED WITH THE EXPECTED NUMBER OF RUNS AMONG THE CHILDREN WHO GUESSED THREE SQUARES

Age (years)		NUMBER OF RUNS					
		5-4 DECKS			7-2 DECKS		
		$r \geq 5$	$r < 5$	Total	$r \geq 5$	$r < 5$	Total
6 + 9	$f(X)$	20.0	11.0	31	23.0	8.0	31
	$E(X)$	9.3	21.7	31	9.3	21.7	31
11	$f(X)$	5.0	16.0	21	2.0	8.0	10
	$E(X)$	6.3	14.7	21	3.0	7.0	10
$P(r)^*$		0.3	0.7	1.0	0.3	0.7	1.0
		$\chi^2_{(6+9)} = 17.6, p < .001, 1 \text{ df}$			$\chi^2_{(6+9)} = 28.8, p < .001, 1 \text{ df}$		
		$\chi^2_{(11)} = 0.35, 1 \text{ df}$					

* $P(r)$ is the probability that such a number of runs will occur by chance (5, p. 57).

Furthermore, the Ss of ages 6 and 9 alternated their guesses significantly more often than would be expected by chance on both kinds of decks whereas the Ss of age 11 did not do so on either kind of deck.³ Table 2 shows these results. The Ss of ages 6 and 9 are treated as a single group because their choices were so similar, as can be seen from Table 1.

These results follow, in large part, from the fact that the Ss of ages 6 and 9 tended to alternate their guesses on each trial, i.e., six runs, whereas the Ss of age 11 did not. Thus, the guesses of the Ss of age 11 were like those of the Ss in Lawlor's study in that they did not alternate regularly whereas the younger Ss did. Only one S of age 11 ever alternated on each successive trial, and he did this only on the 5-4 decks. There were 10 Ss of ages 6 and 9 on the 5-4 decks and 14 Ss on the 7-2 decks who alternated their choices on each successive trial. On both kinds of decks, only 3.1

³ The test is made by comparing the number of Ss whose guesses have five or more runs and the number of Ss whose guesses have less than five runs with the number of Ss in each category that would have been expected by chance.

GERALD GRATCH

Ss would have been expected to act in this way ($Z_{5-4} = 3.11$, $p < .002$; $Z_{7-2} = 4.61$, $p < .0001$).

Lawlor's results suggest the possibility that the maximum run length of children's guesses will tend to be short also. We have already seen from the previous analyses that the *Ss* of ages 6 and 9 tended to guess short runs, their maximum run length being either one (where there are six runs) or two (where there are five runs). The maximum run length of the *Ss* of age 11 was also short, but, unlike the guesses of the younger *Ss*, their maximum run length tended to be exactly two. On the 5-4 decks, 18 of the 21 *Ss* had a longest run of two; three *Ss* had maximum runs of three or one. The probability that the longest run will be exactly two is .60 and, therefore, since 12.6 *Ss* would be expected to guess a maximum run of two by chance, the observed results are significantly different from chance ($\chi^2 = 5.78$, $p < .02$, 1 *df*). The same trend is apparent on the 7-2 decks. Nine of the 10 *Ss* chose so that their maximum run length was two and six would be expected to do so by chance.

In summary, among those *Ss* who guessed *s* three times, the *Ss* of ages 6 and 9 tended to guess alternate events on each successive trial. The *Ss* of age 11 tended to guess that the events would alternate less often, but they did tend to predict that *ss* and *ds* would not occur in runs longer or shorter than two.

Most of the *Ss* who did not guess *s* three times, guessed either two or four *ss*. (The number of *Ss* who guessed zero, one, five, or six *ss* was five on the 5-4 decks and four on the 7-2 decks.) When the *Ss* who guessed some number of *ss* other than three are compared with respect to the number of runs in their guesses, it is found that there are no significant differences between the age groups on either the 5-4 and the 7-2 decks. Also, these *Ss* did not alternate the events they guessed more often than would be expected by chance. However, the longest run in their sequence of guesses also was short, i.e., the length was two. This result is obtained by

TABLE 3

MAXIMUM RUN LENGTH OF *Ss* WHO CHOSE SQUARE TWO OR FOUR TIMES

No. of <i>Ss</i>	MAXIMUM RUN LENGTH					
	5-4 DECKS			7-2 DECKS		
	<i>r</i> = 2	<i>r</i> = 3, 4	Total	<i>r</i> = 2	<i>r</i> = 3, 4	Total
<i>f</i> (<i>X</i>)	19.0	8.0	27	22.0	17.0	39
<i>E</i> (<i>X</i>)	10.8	16.2	27	15.6	23.4	39
<i>P</i> (<i>r</i>)*4	.6	1.0	.4	.6	1.0
	$\chi^2 = 10.38$, $p < .005$, 1 <i>df</i>			$\chi^2 = 4.38$, $p < .05$, 1 <i>df</i>		

* *P*(*r*) is the probability of the maximum run length when there are two or four *ss* and the six events are arranged randomly (4, p. 59).

CHILD DEVELOPMENT

comparing the number of *Ss* whose longest run was two and the number of *Ss* whose longest run was greater than the number expected by chance. Table 3 shows this analysis for the 5-4 and the 7-2 decks.

We are now ready to consider the frequency with which the *Ss* guessed *s* on the two kinds of decks. As one can anticipate from the sequential analyses, there is a significant age trend. The *Ss* of age 11 tended to guess *s* more often on the 7-2 decks than they did on the 5-4 decks, whereas the other two groups did not guess differentially on the two kinds of decks. Tables 4 and 5 show, respectively, the mean number of guesses of *s* of each age group on the two kinds of decks and the results of an analysis of variance of these guesses (11).

TABLE 4

THE MEAN NUMBER OF CHOICES OF SQUARE ON THE 5-4 AND 7-2 DECKS

Age (years)	D E C K		Total
	7-2	5-4	
6	3.25	3.32	3.28
9	3.03	3.21	3.12
11	3.43	2.82	3.12
Total	3.24	3.12	3.18

The scores in the analysis based upon error₁ are correlated, i.e., each *S* had a score based upon the number of times he guessed *s* on the 5-4 decks and another score based upon the number of times that he guessed *s* on the 7-2 decks. There is a significant interaction between deck and age. To evaluate the difference between the six means, a multiple range test was used (4, 9). Of the 15 possible pairs of comparisons, two were significant at the 5 per cent level: (a), the *Ss* of age 11 guessed *s* significantly more

TABLE 5

ANALYSIS OF VARIANCE OF CHOICES OF SQUARE

Source	Sum of Squares	df	Mean Square	F
Decks	0.60	1	0.60	1.02
Age Groups \times Decks	5.08	2	2.04	3.50*
Error ₁	47.32	81	0.58	
Age Groups	0.96	2	0.48	
Between <i>Ss</i> (error ₂)	52.68	81	0.65	
Total	106.64	167		

* Significant at the 5 per cent level.

GERALD GRATCH

often on the 7-2 decks than they did on the 5-4 decks; and (b), on the 5-4 decks, the *Ss* of age 6 guessed *s* significantly more often than the *Ss* of age 11. In the analysis based upon the "Between *Ss*" error term (error_2), each *Ss*' score is the sum of the number of times he chose *s* on both kinds of decks. There are no significant differences in this latter analysis.

Before discussing these results, we should note that the *Ss* tended to guess *s* three times to a striking degree (as will be shown below). In the cases of the *Ss* of ages 6 and 9, this "peakedness" in the distribution of their guesses of *s* follows directly from the fact that they tended to alternate their guesses of *s* and *d* on each successive trial. However, the *Ss* of age 11 did not alternate more often than would be expected by chance, and yet, on the 5-4 decks, many of them guessed *s* three times. The possible implications of the behavior of the *Ss* of age 11 will be discussed below. Table 6 shows that more *Ss* did guess *s* three times than would be expected by

TABLE 6
THE NUMBER OF CHILDREN CHOOSING SQUARE THREE OR SOME
OTHER NUMBER OF TIMES

Age (years)	NUMBER OF SQUARES					
	5-4 DECKS			7-2 DECKS		
	<i>s</i> = 3	<i>s</i> ≠ 3	Total	<i>s</i> = 3	<i>s</i> ≠ 3	Total
6	17	11	28	15	13	28
9	14	14	28	16	12	28
11	21	7	28	10	18	28
Total	52	32	84	41	43	84
$P_1(s)^*$31	.69	1.00	.31	.69	1.00
$P_2(s)^*$30	.70	1.00	.10	.90	1.00

* $P_1(s)$ is the probability that a subject will choose square three times if he is choosing at random and is choosing *s* with a probability of one-half. $P_2(s)$ is the probability that a subject will choose square three times if he is choosing randomly and is choosing *s* with a probability equal to the probability of occurrence of square in the decks.

chance on either the assumption that they were guessing *s* and *d* equally often [$P(s_{7-2}) = P(s_{5-4}) = 0.5$] or on the assumption that they were guessing *s* and *d* with the same probability with which they would tend to occur in the decks [$P(s_{7-2}) = 0.78$; $P(s_{5-4}) = 0.56$].

With respect to the first assumption, that the *Ss* were guessing *s* and *d* with the same probability, significantly more *Ss* in all three age groups guessed *s* three times than would be expected by chance on the 5-4 decks. On the 7-2 decks, with the same assumption, more *Ss* of ages 6 and 9 guessed *s* three times than would be expected by chance. But no more of the *Ss* of age 11 guessed *s* three times than would be expected by chance,

CHILD DEVELOPMENT

and the majority of the *Ss* of age 11 chose *s* four times on the 7-2 decks. On the other hand, under the second assumption, that the *Ss* were guessing *s* with the same probability with which it would occur in the two kinds of decks, more *Ss* in all three age groups guessed *s* three times than would be expected by chance on both kinds of decks.

In summary, all the *Ss* tended to guess so that their runs were short. However, the *Ss* of ages 6 and 9 tended to alternate the events that they guessed on each successive trial on both kinds of decks. The *Ss* of age 11 alternated their guesses of events less often than the younger *Ss*. Secondly, the frequency of the latter *Ss*' guesses of *s* and *d* tended to match the probabilities of *s* and *d* in the respective decks whereas the younger *Ss* did not guess differentially on the two kinds of decks.

Finally, it should be noted that the *Ss* did not seem to have any special preference for either the *s* or the *d* alternatives. If the *Ss* as a group had "liked" one of the two choice-alternatives more, they would have manifested their preference on the first trial, for that deck had almost the same number of *ss* and *ds*, i.e., 5-4. On this first trial, the guesses of the three groups did not differ significantly from one another and there was no significant tendency for the *Ss* to guess *s* more often than *d*. Moreover, the *Ss* did not seem to display any special concern over the markings of the cards and they did not tend to say that they "liked" either *ss* or *ds* more.

DISCUSSION AND CONCLUSIONS

As a group, the *Ss* of age 11 seem to have followed the "plan" of trying to guess as many *ss* and *ds* as there were in the two kinds of decks. Indeed, some of them explained their guesses in just that way. Of course, the *Ss*' explanations showed that they were confusing sampling with replacement (the conditions of this experiment) with sampling without replacement (their explanation). That they were following the plan outlined above is supported by the distribution of their guesses. Considering the restrictions placed upon their guesses, namely, they had to make only six guesses on each kind of a deck and they had a tendency to choose *s* not more than twice in succession, we would expect the pattern observed: three *ss* on the 5-4 decks; four *ss* on the 7-2 decks. However, this explanation is clearly an incomplete account of the *Ss*' behavior, for it does not clarify their tendency to have a maximum run length of two. The author can offer no explanation for this sequential tendency at the present time. He only can emphasize the fact that the *Ss* of age 11 seem to have had the beliefs about random events that Lawlor and others have found characteristic of adults.

The remarks of the younger *Ss* also suggest that, as a group, these *Ss* thought they knew what was going to happen, and their comments provide a possible explanation of their behavior. Many of the *Ss* of ages 6 and 9, particularly of age 6, explained that they chose in the way that they did because it was "fair." They would say such things as "the *ss* and *ds* ought

to take turns," "the *ss* and *ds* take turns," etc., and we have already seen that they tended to alternate the events in this way, irrespective of the kinds of decks. Thus, it is possible that they were acting upon the basis of an expectation that the events would conform to a moral order, e.g., Piaget's view (12) that young children's thinking is "morally realistic" in the sense that they tend not to distinguish between moral rules and physical laws and to project the former upon physical events. This explanation is consistent with the fact that the older children had relatively more experience with numbers and the younger children were probably relatively more concerned with the problem of learning what is right and what is wrong. However, it is possible that their moralizing simply may have been their way of rationalizing what they did for quite different reasons (*see below*).

In support of the view that the *Ss*' guesses indicated how they thought the events would occur, it should be noted first that the *Ss* in all three groups gave every indication that they were expecting their guesses to match a set of events. The *Ss* frequently asked, both throughout the guessing sequence and at the end of it, about the number of guesses that they would get "right" and whether what they guessed would happen.

There are several general theories of learning and performance which deal with alternation of responses in a way which is consistent with the behavior of the *Ss* of ages 6 and 9. Walker's "reaction decrement" hypothesis (16), Hull's "reactive inhibition" assumption (5), and the Gestalt hypothesis that children and adults have a preference for balanced and repetitive actions and events (e.g., Koffka, 8) all predict that *Ss* will alternate binary responses on pairs of successive trials. These latter views would permit an interpretation of the *Ss*' behavior in less vague and "mentalistic" terms than have been used in the discussion above. Unfortunately, these views do not attempt to explain what will happen when *Ss* make longer sequences of choices. Further, they do not seem to afford any simple basis for understanding why the *Ss* of age 11 made different patterns of choices than the younger *Ss*.

Therefore, the principal conclusions of this study must be empirical rather than explanatory. Some developmental facts have been presented, and they do not appear to be isolated phenomena. For example, other studies have found that young children (1, 13) and adults (7, 10) tend to predict that binary events will alternate on successive trials. Secondly, these studies have utilized a variety of situations, e.g., which of two teams will win the game, which of two nonsense syllables will appear, which hole will the marble roll into. Further, these kinds of facts are relevant to a variety of choice situations in which *Ss* make sequences of choices, e.g., psychophysical experiments (e.g., 14), econometric studies, learning studies. Therefore, there would seem to be a need for further exploration of the sequences of choices of adults and children and for an incorporation of these facts into present theories of learning and performance, which for the most part have not attempted to account for such response contingencies, e.g., Estes (3).

CHILD DEVELOPMENT

SUMMARY

Children 6, 9, and 11 years old predicted what the top card would be of a sequence of decks of cards. The first six decks were composed of five cards marked with a blue square and four cards marked with a blue diamond. The next six decks each had seven squares and two diamonds. The children were told what the composition of each deck was. After they made all 12 guesses, they were shown what the top cards were. The Ss of ages 6 and 9 tended to alternate their choices on each successive trial on both kinds of decks. It was suggested that their behavior might have been a result of a belief on their part that the events should or would be "fair" and should or would "take turns." The Ss of age 11 did not tend to alternate their choices on each successive trial, but they did predict that *ss* and *ds* would never occur more than twice in succession. They responded to the cue of deck composition by tending to predict that *s* would occur three times on the 5-4 decks and four times on the 7-2 decks. It was suggested that their behavior could be explained in terms of their having made an attempt to match, with their predictions, the probability of *s* in the two kinds of decks. The relation between the behavior of the Ss of age 11 and adult Ss was discussed.

REFERENCES

1. COHEN, J., & HANSEL, C. E. M. The idea of independence. *Brit. J. Psychol.*, 1955, 46, 178-190.
2. EDWARDS, W. The theory of decision making. *Psychol. Bull.*, 1954, 51, 380-417.
3. ESTES, W. K., & BURKE, C. J. A theory of stimulus variability in learning. *Psychol. Rev.*, 1953, 60, 276-286.
4. DUNCAN, D. B. Multiple range and multiple F tests. *Biometrics*, 1955, 11, 1-42.
5. FELLER, W. *An introduction to probability theory and its applications*. New York: Wiley, 1950.
6. HULL, C. L. *Principles of behavior*. New York: D. Appleton-Century, 1943.
7. JARVICK, M. H. Probability learning and a negative-recency effect in the serial anticipation of alternating symbols. *J. exp. Psychol.*, 1951, 41, 291-297.
8. KOFFKA, K. *The growth of the mind*. New York: Harcourt, Brace, 1924.
9. KRAMER, C. Y. Extension of multiple range tests to group correlated adjusted means. *Biometrics*, 1957, 13, 13-18.
10. LAWLOR, W. Subjective probability in sequential uncertainty situations. Unpublished doctoral dissertation, Univer. of Chicago, 1956.

GERALD GRATCH

11. LINDQUIST, E. L. *Design and analysis of experiments in psychology and education*. Boston: Houghton Mifflin, 1953.
12. PIAGET, J. *The child's conception of the world*. London: Routledge & Kegan Paul, 1951.
13. PIAGET, J., & INHELDER, B. *La genese de l'idee de hasard chez l'enfant*. Paris: Presses Univer. De France, 1951.
14. SENDERS, V. L., & SOWARDS, A. Analysis of response sequences in the setting of a psychophysical experiment. *Amer. J. Psychol.*, 1952, 65, 358-374.
15. SKINNER, B. F. The processes involved in the repeated guessing of alternatives. *J. exp. Psychol.*, 1942, 30, 495-502.
16. WALKER, E. L. The duration and course of reaction decrement and the influence of reward. *J. comp. physiol. Psychol.*, 1956, 49, 167-175.

Section C

The following is a list of the names of the persons who have been elected to the office of the President of the United States since the year 1789. The names are arranged in alphabetical order of the year in which they were elected.

George Washington 1789
John Adams 1793
Thomas Jefferson 1801
James Madison 1809
James Monroe 1817
John Quincy Adams 1825
Andrew Jackson 1829
Martin Van Buren 1837
William Henry Harrison 1841
John Tyler 1845
Zachary Taylor 1850
Franklin Pierce 1853
Abraham Lincoln 1861
Andrew Johnson 1865
Ulysses S. Grant 1869
Rutherford B. Hayes 1877
James A. Garfield 1881
Chester A. Arthur 1881
Grover Cleveland 1885
Benjamin Harrison 1889
William McKinley 1897
Theodore Roosevelt 1901
William Howard Taft 1909
Woodrow Wilson 1913
Warren G. Harding 1921
Calvin Coolidge 1925
Herbert Hoover 1929
Franklin D. Roosevelt 1933
Dwight D. Eisenhower 1953
John F. Kennedy 1961
Lyndon B. Johnson 1963
Richard M. Nixon 1969
Jimmy Carter 1977
Ronald Reagan 1981
George H. W. Bush 1989
Bill Clinton 1993
George W. Bush 2001
Barack Obama 2009
Donald Trump 2017

THE EFFECT OF AGE AND RELATIVE BRIGHTNESS ON ASSOCIATIVE LEARNING IN CHILDREN

DONALD AKUTAGAWA*^{1,2} and E. PAUL BENOIT*¹

Governor Bacon Health Center

Developmental theorists and child psychologists have placed considerable emphasis on the influence of "maturational" processes on psychological and behavioral development. However, during recent years there have been relatively few systematic investigations concerned with this problem. A major difficulty, of course, has been the lack of clear theoretical formulations that specify the way in which maturational and environmental factors combine to produce effects on psychological development. Hebb's theory (2) is a step in the direction of developing such a structure. Previous investigations of age changes, as well as age differences, in learning have yielded conflicting results. Munn (5) and Razran (6), in the light of these disagreements, have pointed out, however, that these studies differ on a number of important variables other than age. Practice, motivational conditions, and type and difficulty level of tasks vary so widely that generalizations concerning the age variable are difficult to make. Nevertheless, there is abundant evidence that maturational factors affect learning (5).

The present experiment was designed to determine the relative contributions of chronological age (maturation index) and IQ (relative brightness or intelligence index) in tasks of simple associative learning.

According to Hebb (2), through repeated simultaneous or sequential activation, neurons are organized into functional networks or "cell assemblies" capable of brief closed circuit activity after the cessation of stimulation. Such "cell assemblies" through repeated interaction with other similarly derived assemblies become organized into more complex systems, "phase sequences," which are capable of even more extended poststimulus activity. A still higher order organization of "phase sequences" in a series constitutes a "phase cycle." Even higher order "cycles of cycles" may be formed through recurrent stimulation. These complex systems form the basis for conceptual processes (thinking) and are considered to be relatively permanent structures.

As indicated above, the formation and development of these neural organizations is viewed by Hebb as a continuing one. Frequent activation

* Governor Bacon Health Center, Delaware City, Delaware.

¹ The writers are indebted to Dr. Charles D. Smock, University of Pennsylvania, for invaluable consultation and assistance in the analysis of data and for helpful criticism of this report.

² Also at the Child Study Center, Philadelphia.

CHILD DEVELOPMENT

of such cell systems increases the ease of arousal of each system as well as the mutual activation of several systems. Such development progresses with increasing age. The limit to the complexity of organization is presumably set by the structural arrangement and the number of neural cells. The degree of complexity within this limit, however, is a function of experiential factors. The innate factors presumably set the upper limit to the level of complexity possible, but whether that limit is closely or scarcely approximated is dependent upon the patterns of activation which occur during the individual's lifetime.

Hebb equates the degree of complexity attained by the neural systems with the usual meaning of intelligence, "the average level of performance or comprehension." The innately determined limit is called by Hebb the "innate potential." While the latter cannot be determined, the former is subject to estimation by means of conventional intelligence testing methods.

For the purpose of the following discussion, let simple concepts like "chair," "table," "cat," etc., be represented by a *phase sequence*. Broader, more inclusive concepts such as "furniture" and "animal" can be represented at the neurological level as *phase cycles*.

According to Hebb (2) a class of items may be considered as a "concept" or "conceptual type," of which individual members are specific instances. A concept would presumably involve one or more "phase sequences" depending upon its complexity. The ease of activating different phase sequences or systems of phase sequences would depend upon the number of subsystems they had in common, which in turn would depend upon the number of times the subsystems had been jointly interactive. In addition, higher order, more complex phase sequences or "phase cycles" or series of cycles could be formed. For example, "chair" and "table" as relatively simple concepts would through repetition come to have some subsystems or "cell assemblies" in common. With further development they would become part of a more inclusive phase cycle "furniture," including both phase sequences. This last concept would be formed according to Hebb's analysis simply by seeing the two kinds of objects together frequently, or by hearing the two together frequently, etc. For children of school age and for adults, associating such concepts should be relatively easy since they are familiar aspects of experience and are frequently experienced together. Other such pairs of objects are "banana-apple," "boy-girl," "cat-dog," etc., within the cycles, "fruit," "sex differences," and "animal," respectively.

A more difficult task would be to learn the associations between quite different concepts, such as "clock" and "elephant," or "star" and "carrot." In Hebb's framework such diverse concepts might be represented at the neurological level by quite different phase sequences which would be subsystems within quite different phase cycles. For example, a phase sequence corresponding to the concept "elephant" could be part of a phase cycle "animal," or a cycle of "zoo," or of "big." But the phase sequence corresponding to "clock" would not be likely to be a part of the same cycles, or at least not a frequently activated part of them. Rather, it would be a part of

cycles like "time," "timepiece," or even "furniture." This kind of learning task would require the joint activation of different *phase cycles* rather than the joint activation of *different phase sequences within a single phase cycle*, as would be the case in the familiar concept pairs, "chair-table," "boy-girl," etc.

A still more difficult task would be to activate two different phase sequences from different phase cycles in a situation in which the tendency to activate only one of the phase cycles is dominant. For example, if the stimulus "chair" has been recently associated with "table" within the phase cycle "furniture," and then "chair" is paired with "pig," a member of a different cycle, the dominant tendency should be to activate the two phase sequences within the cycle "furniture" rather than the two sequences from two different cycles. In order to overcome this tendency, some mechanism would have to be involved to inhibit the first phase cycle to allow the joint activation of the two new phase sequences.

These considerations allow for the construction of learning tasks of three levels of difficulty. At the first level, familiar objects commonly experienced together would be paired. At the second level, concepts from very different higher-order categories or phase cycles would be arbitrarily paired. For the third level, one of the pairs used in the first level could be arbitrarily paired with a very different object.

As indicated above, a task of simple associative learning such as rote learning of paired pictures involves at the neurological level the establishment of functional interaction between cell assemblies or phase sequences rather than between complex phase cycles. The ease of activation of structures is a function of their degree of organization which is in turn a function of exercise. Since exercise is largely limited by time, older persons should master simple associative learning tasks more quickly than younger persons because the relevant neural circuits are better established and are therefore more easily and quickly activated. If Hebb's interpretation of measured intelligence level as the established degree of complexity of neural organization is accurate, IQ should have little relation to performance in simple associative learning.

Since the learning tasks chosen for this investigation are of relatively simple associative learning, efficiency of learning should be more related to the age variable than to measured intelligence. Hypothesis 1 may be stated: *Performance level on all three tasks used in this study should be more related to chronological age than to IQ.*

On the easiest task of associating familiar pairs of objects, the individual concepts should be well established in the particular population used in this study. Second order relationships between the pairs should be better formed in older persons and permit of more efficient associative learning. Intelligence or relative brightness might reasonably be assumed to be an important factor in forming an associative connection between the two objects, but Hebb's formulation leads to the deduction that it would be of very little consequence in such simple associative learning. In the light

CHILD DEVELOPMENT

of this viewpoint, hypothesis 2 may be stated: *Performance level on a task involving the association of two concepts which are members of a single second order category should be different between the age groups in favor of the older group, but not different between the two levels of intelligence.*

On the task of intermediate difficulty, associating familiar pairs of objects from two very different higher order categories, not only should the phase sequences represented by the individual first order concepts be well established, but multiple connections should be formed with other phase sequences from other such additional connections and should by chance have more subsystems in common between paired concepts, thus facilitating associative learning. In accordance with Hebb's viewpoint, since this task is more difficult than the former one, and since more complex phase cycles should be involved, intelligence may be more relevant to good performance on it. A third hypothesis may be stated: *Performance level on a task involving the association of two concepts which are members of different second order categories should be different between the age groups in favor of the older group, and different between the intelligence groups in favor of the brighter.*

On the most difficult of the chosen tasks, the same considerations as for the first two apply. But in addition, since this task should be more complex than the other two, relative brightness should play an even more influential role. Also, because of the interference factor, two different consequences might be deduced. In the light of the considerations presented above, older persons should have more functional associative connections with already acquired concepts and may be expected on this basis to show more difficulty in responding with the correct response than the younger persons who have fewer associative connections which are also less well developed. However, on this task, interfering associative circuits within a single cycle must be inhibited in order for activation of two phase sequences from two different phase cycles to take place. Older persons have acquired many nonspecific functions such as inhibition, mnemonic devices, etc., which though present in the younger persons also, should be less well learned by the latter, and therefore less effective in operation. A fourth hypothesis based on these considerations is identical with the previous one, and a hypothesis combining both may be stated as hypothesis 3: *Performance level on tasks involving the association of two concepts which are members of different second order categories should be different between the age groups in favor of the older group, and different between the intelligence groups in favor of the brighter.*

METHOD

Subjects

Forty-one boys, both Negro and white, ages 8 to 13 inclusive, in residence at the Governor Bacon Health Center in Delaware City, Delaware, were selected as subjects for this study. On the basis of chronological age

and IQ ratings from routine psychological examinations, they were assigned to one of four groups: Younger, Subaverage—composed of 9 children 8 to 10 years of age, and within the IQ range 70 to 89; Younger, Average—composed of 9 children 8 to 10 years of age, and within the IQ range 90 to 110; Older, Subaverage—composed of 11 children 11 to 13 years of age, and within the IQ range 70 to 89; Older, Average—composed of 12 children 11 to 13 years of age, and within the IQ range 90 to 110.

The Learning Tasks

The selection of learning tasks was based upon Hebbian theory as described above. An essential modification of Ruch's (7) verbal learning tasks designed to provide tasks differing in difficulty, novelty, and the degree to which prior experience would facilitate or disrupt present learning was utilized. Pictures of common objects were chosen from a popular children's book as stimulus and response items. Three lists of paired pictures were prepared:

Task I, Familiar Associates—paired pictures of common objects frequently paired together environmentally and linguistically (e.g., Chair-Table, Cat-Dog, etc.)

Task II, New Associates—paired pictures of common objects, but ones usually not contiguous in time or space, or linguistically (e.g., Clock-Elephant, Star-Carrot, etc.)

Task III, Interference Associates—paired pictures of common objects of which the first was identical to the stimulus pictures of the *Familiar Associates*, but the response item of which is different and not usually associated with the first (e.g., Chair-Pig, Cat-Bed, etc.)

Each task included eight pairs of pictures. Table 1 presents the list of pictures for each of the three tasks.

TABLE 1
PICTURES USED IN THE THREE LEARNING TASKS

<i>Familiar Associates</i>		<i>Arbitrary Associates</i>		<i>Interference Associates</i>	
Horse	— Cow	Tricycle	— Monkey	Horse	— Baby
Shoes	— Stockings	Rooster	— Church	Shoes	— Corn
Boy	— Girl	Wagon	— Owl	Boy	— Lion
Hen	— Eggs	Bread	— Rake	Hen	— Dress
Banana	— Apple	Star	— Carrot	Banana	— Airplane
Car	— Truck	House	— Bird	Car	— Goose
Cat	— Dog	Man	— Ship	Cat	— Bed
Chair	— Table	Clock	— Elephant	Chair	— Pig

Procedures

The instructions for this study were a simplified version of those used by Ruch. A short demonstration of the tasks and an introductory trial with

CHILD DEVELOPMENT

the experimental list was given to each subject. All subjects were able to name the object class of all the items on the lists. The lists of paired associates were presented by the anticipation method, and correction procedure was utilized.

The stimulus items were exposed for four seconds followed by an additional four seconds of exposure to both the stimulus and response items. Four seconds were allowed between each stimulus-response pair and one minute between each of the three tasks. The three tasks were presented in order, I, II, III. Each task of eight pairs was presented six times to each subject for a total of 48 trials. The order of the items was changed after each trial to prevent serial learning.

TABLE 2
ANALYSIS OF VARIANCE

Source of Variation	Sum of Squares	df	Mean Square	F
Between CA	1,613.33	1	1,613.33	19.72**
Between IQ	163.33	1	163.33	1.996
Interaction: CA \times IQ	40.84	1	40.84	.50
Residual between groups	2,945.13	36	81.81	
Total between	4,762.63	39		
Within lists	10,186.52	2	5,093.26	224.47**
Interaction: Tasks \times CA	195.82	2	97.91	4.32*
Interaction: Tasks \times IQ	7.22	2	3.61	
Interaction: Tasks \times CA \times IQ	24.11	2	12.05	
Residual within groups	1,633.70	72	22.69	
Total within	12,047.37	80		
Total	16,809.97	119		

* Significant beyond .05 level.

** Significant beyond .01 level.

RESULTS AND DISCUSSION

Table 2 presents the analysis of variance.³ The mean scores for each of the three tests in the two age groups are presented in Table 3. Each score unit in this table and the following one represents one correct anticipation of a pair of pictures. Since there were eight pairs of pictures in each task, and each subject was given six trials with each list of picture pairs, a maximum of 48 points could be earned on each of the three tasks. In Table 4 are presented the mean learning scores on each of the three tasks for the two intelligence level groups.

³ In order to facilitate computation, two of the four groups were increased to 10 by adding mean scores for the particular group; the other two were decreased to 10 by random selection and elimination.

TABLE 3
MEAN LEARNING SCORES FOR TWO AGE GROUPS

Group	N	Mean C.A.	Familiar Associates	New Associates	Interference Associates
Younger	18	9.44	41.33	24.56	16.16
Older	23	11.95	45.43	30.91	25.43
Diff $M_0 - M_1$..			4.10	6.35	9.27
$t =$			2.63*	3.32**	3.32**

* Significant beyond .02 level.

** Significant beyond .01 level.

Hypothesis 1 is confirmed. Only the variance between the age groups, the within variance for tasks, and their interaction are significant. Performance of the tasks appears to be clearly more related to CA than to IQ. The large variance within tasks is a function of the different levels of difficulty which they present to the subject as planned for this study.

Hypothesis 2 is confirmed. The difference between the two age groups as shown in Table 3 is significant, but the difference between the two intelligence groups is not, as shown in Table 4.

TABLE 4
MEAN LEARNING SCORES FOR TWO INTELLIGENCE GROUPS

Group	N	Mean IQ	Familiar Associates	New Associates	Interference Associates
Subaverage	20	84.5	42.65	26.95	19.50
Average	21	96.3	44.57	29.23	23.14
Diff $M_0 - M_1$..			1.92	2.28	3.64
$t =$			1.23	.98	1.19

Hypothesis 3 is partially confirmed. The difference between the two age groups is significant, but the difference between the two intelligence groups is not.

Older subjects clearly perform better on each task than do the younger ones. The supposedly reasonable expectation that efficiency of learning is a simple function of measured intelligence level is not borne out. It appears that, in conformity with deductions from Hebbian theory, that lower levels of associative learning are more determined by gross amount of previous experience than by relative intelligence. This is true even with highly overlearned associations such as those of Task I. Table 5 presents the mean scores for the yearly age levels. A fairly consistent trend increasing with age appears even here.

CHILD DEVELOPMENT

TABLE 5

MEAN LEARNING SCORES FOR CHRONOLOGICAL AGE GROUPS

CA	N	Mean IQ	Familiar Associates	Arbitrary Associates	Interference Associates
8	6	85.67	36.33	21.50	13.50
9	7	88.40	43.14	26.14	16.86
10	5	88.2	44.80	26.00	18.40
11	6	79.5	43.83	28.50	22.50
12	13	92.6	45.62	30.92	26.69
13	4	93.5	47.25	36.75	25.75

Apparently the tasks chosen for this study are of such simple associative learning that relative brightness is not the most important factor in proficiency. It is evident from the trends present, however, that relative brightness should become much more crucial with more complex learning tasks than these. A further study to determine whether this is so, and also at what degree of task complexity relative brightness becomes more crucial than age-concomitant learnings, is planned.

Analysis of the types of errors made shows some interesting relationships. Table 6 presents this information. Three major classes of errors are tabled, *No Response*, *Random Response* (response not on the lists), and *Correct Elsewhere* (on the same or other lists used in this study).

From the patterns of errors on the different tasks, perhaps some evidence may be adduced regarding an inhibition factor. If no inhibition factor were involved, greater numbers of alternate associative connections (*a la* Hull), or greater numbers of subsystems of phase sequences (*a la* Hebb) should lead to particular kinds of errors more frequently than others. In particular, on task II, there should be more errors in the older group resulting from responses that are correct on task I (rows 2 and 5, column 4); on task III, there should be more errors in the older group resulting from responses that are correct on task I and task II (rows 3 and 6, columns 4 and 5); and the older group should make greater numbers of random response errors (rows 1 through 6, column 3). However, if an inhibition factor were involved, suppressing incorrect responses, the opposite findings should be the case. Although the number of erroneous transfers from tasks I and II are small, all four of the trends are in the direction compatible with an inhibition factor in the older groups.

Another interesting finding is the fact that the younger group made more errors than the older group on tasks I and III, but fewer errors on task II. This is consistent with the viewpoint that older persons would have more subsystems in common between well established concepts, and that selection from among them would be more difficult where inhibition is not strongly involved. On task III where the conflict between competing re-

DONALD AKUTAGAWA and E. PAUL BENOIT

TABLE 6

ANALYSIS OF TYPES OF ERRORS

Task	No Response	Random Response	CORRECT ELSEWHERE			Total
			Error but Correct for:			
			Task I	Task II	Task III	
<i>Older Group</i>						
I	55	0	4	0	0	59
II	271	0	1	123	0	395
III	327	1	16	7	186	537
<i>Younger Group</i>						
I	112	2*	5	0	0	119
II	177	3	6	122	0	308
III	375	5	19	14	172	585
<i>Average</i>						
I	69	0	2	0	0	71
II	143	2	3	137	0	285
III	340	4	8	19	154	525
<i>Subaverage</i>						
I	98	2*	7	0	0	107
II	305	1	4	108	0	418
III	344	2	27	2	204	579

* Responses actually correct for task II, but classed as random response because subject had no experience with task II pictures when error was made.

sponses is heightened, inhibition should be much more strongly called into action, allowing for more efficient response. The smaller number of total errors on task III would support this hypothesis.

Older subjects clearly perform better on each task than do the younger ones. Learning does not appear to be a simple function of measured intelligence level. These findings have many implications for the education of children of lower intellectual ability levels as indicated by Benoit (1). First, it appears in conformity with deductions from Hebbian theory that lower levels of associative learning are more determined by gross amount of non-specific previous experience than by relative intelligence. More optimism about the educability of the mentally retarded may be warranted. Tasks requiring simple associative learning rather than higher level abstract and conceptual learnings may be acquired almost as well by the mentally retarded as by persons of higher relative intelligence. Secondly, the crucial importance of early learning for later learning is strongly emphasized by Hebb's formulations and supported by these findings. Carefully planned, systematic learning experiences progressing gradually from simple to more complex levels is probably necessary for the most efficient learning. For maximum efficiency of learning, such planning will necessarily include

CHILD DEVELOPMENT

analysis of the component, antecedent learnings necessary for particular later learning, and their relative temporal occurrence in relation to the others. Thirdly, under such a program of highly structured learning, relatively high levels of proficiency in very specific, selected skills may be attained by persons of limited intellectual endowment.

SUMMARY

This study was concerned with the comparative efficiency with which children of different age and measured intelligence levels learn tasks of differing degrees of difficulty. Deductions from Hebb's formulations concerning the relative importance of age and intelligence were tested. Four different groups of boys of different age and IQ levels were used in this study, performing paired associate tasks presented by the anticipation method with correction. The results of the investigation fully supported the deductions in two cases and partially confirmed two others. A gradual, consistent, and statistically significant increase in efficiency of performance with age in simple associative learning tasks was found. Although the trends were somewhat similar in the case of learning efficiency and intelligence, the results were not significant. The expectation that learning efficiency is a simple function of measured intelligence was not borne out. These findings are interpreted as supporting Hebb's theoretical formulation.

REFERENCES

1. BENOIT, E. P. Relevance of Hebb's theory of the organization of behavior to educational research on the mentally retarded. *Amer. J. ment. Def.*, 1957, 61, 497-507.
2. HEBB, D. O. *The organization of behavior*. New York: Wiley, 1949.
3. KORCHIN, S. J., & BASOWITZ, H. Age differences in verbal learning. *J. abnorm. soc. Psychol.*, 1957, 54, 64-69.
4. LINDQUIST, E. L. *Design and analysis of experiments in psychology and education*. Boston: Houghton Mifflin, 1953.
5. MUNN, N. L. Learning in children. In L. Carmichael (Ed.), *Manual of child psychology*. (2nd Ed.) New York: Wiley, 1954. Pp. 374-458.
6. RAZRAN, G. H. S. Conditioned responses: an experimental study and a theoretical analysis. *Arch. Psychol.*, NY, 1935, No. 191.
7. RUCH, F. L. The differentiative effects of age upon human learning. *J. gen. Psychol.*, 1934, 11, 261-286.

ROLE-TAKING IN CHILDHOOD AND ITS CONSEQUENCES FOR SOCIAL LEARNING¹

ELEANOR E. MACCOBY*

Harvard University

Let us assume that a large part of what a young child learns is acquired through the process of instrumental learning. This process involves selection of some of the child's overt actions via reward, and the elimination of others either through nonreward or through punishment. I wish to propose that concurrently the young child is acquiring a repertoire of behavior through a different process: that of practicing covertly the characteristic actions of other people with whom he interacts.

Since reproducing other people's actions has been labeled "imitation," and imitation has been described as a product of instrumental learning, [cf. Miller and Dollard (3)], it may be well to begin with a detailed description of the kind of imitation I am talking about, and to consider whether it does indeed follow the principles of instrumental learning.

Let us draw on some of the ingenious and marvelously detailed descriptions of imitation in the first year of life reported by Piaget (6). In describing the behavior of one of his infant daughters, Piaget notes that at first she seemed to react to voices without any apparent effort to imitate specific sounds. Here is a report when the infant was 1 month and 25 days old:

She was watching me while I said "a, ha, ha, ra" etc. I noticed certain movements of her mouth, movements not of suction but of vocalization. She succeeded once or twice in producing some rather vague sounds, and although there was no imitation in the strict sense, there was obviously vocal contagion (6, p. 10).

Genuine imitation began at about 3 months of age. Piaget describes the sequence as follows:

I noted a differentiation in the sounds of her laughter. I imitated them. She reacted by reproducing them quite clearly, but only when she had uttered them immediately before (6, p. 10).

* Laboratory of Human Development, Harvard University, Palfrey House, Cambridge 38, Massachusetts.

¹ Revised from a paper delivered at a Colloquium at Duke University in March, 1958.

CHILD DEVELOPMENT

Piaget uses the term "circular reaction" to describe this kind of imitation, and indeed we are reminded immediately of Holt's circular reflex principle. Piaget specifically disavows the interpretation that simple conditioning is the explanation of what he has observed. He says:

If we consider first of all the circular reaction, how are we to explain the fact that perception "combines" with a movement and acquires a motor power which is not inherent in it? It cannot be a question of passive association, otherwise the child's activity would be the plaything of the most haphazard occurrences (6, p. 17).

If we were to rephrase the problem in language more familiar to American psychologists, the question would be "How are we to explain the fact that a perceived stimulus (such as the sound "ba") comes to elicit a response which it did not formerly elicit?" The stimulus is the action of another person, and the response is a similar, or almost identical, response. Piaget's disavowal notwithstanding, I believe his descriptions are consonant with the interpretation that simple conditioning (or what he calls "passive association") is indeed occurring. We note that in order for imitation of a sound to occur the child must first have made it spontaneously. According to the circular reflex principle, the child's utterance stimulates his ear at the same moment that the motor activity required to produce the sound is occurring; thus, by the Guthrie principle of association by contiguity, the auditory stimulus comes to have the power to elicit this particular response of the vocal apparatus, and we have the conditions laid down for repetitive babbling: that is, the child's hearing his own voice saying "ba" will stimulate him to say "ba" again. If someone else makes a sound similar enough to "ba," the child will also be stimulated to say this syllable, and we have imitation. It should be evident at once that it should be easier to establish imitation of speech sounds than of bodily movements and especially of facial expressions. For the child always stimulates his own ear when he vocalizes, and the stimulus is much the same as that which occurs when others vocalize the same phoneme in his presence. But he makes certain movements without seeing them (as for example, when he smiles). True, he has the motor feed-back from his response. But the feeling of his facial muscles moving is quite a different stimulus from the sight of someone else's mouth smiling. There is no reason why the sight of someone else wrinkling his nose, for example, should stimulate the child to do the same, since this stimulus is not ordinarily present when he wrinkles his own nose. Piaget gives us some descriptions bearing on this point. He had been trying for some time to get one of his infant daughters to imitate various hand movements that he made. When the child was nearly seven months old, he would slowly open and close his hand in front of her, and she would not imitate this movement. Then he reports the following when the child was 8½ months old:

I observed that she alternately opened and closed her right hand, watching it with great attention as if this movement, as an isolated schema, was new

ELEANOR E. MACCOBY

for her. I made no experiment at that point, but the same evening I showed her my hand as I opened and closed it rhythmically. She thereupon imitated the movement, rather awkwardly but quite distinctly. She was lying on her stomach and not looking at her hand, but there was a clear correlation between her movements and mine (6, p. 23).

Here we have a situation in which, once the child makes a movement that she can see, the sight of her own movement seems to become an adequate stimulus for eliciting the movement. Only after this learning has occurred can the sight of someone *else* making the movement elicit the response. Similarly, Piaget reports trying to produce imitation of a kind of clapping movement, and being unable to get the response until he observed the child (now 6 months old) spontaneously bringing her hands together and moving them apart, at which point he moved his hands in front of her in just the way her own hands had previously moved, and she imitated him. Piaget's own conclusion is that an infant of this age can imitate any hand movement that he can make spontaneously, but cannot imitate new movements.

What about movements that the child cannot see, such as facial expressions? Piaget found it much more difficult to obtain imitation of such movements. Even though the child spontaneously put out its tongue, or wrinkled its nose, for example, it would not perform these actions when someone else did them. What proved to be necessary was what Piaget called "training in imitation"; that is, he had to wait till the child made the movement spontaneously, then at once make the response himself to establish a connection between the stimulus of his facial movement and the child's matched response. In other words, he had to imitate the child, before the child could imitate him, when the movement was one the child could not see himself make. Presumably, a child could learn to imitate such movements also by watching his own face in a mirror.

While it is likely that Piaget showed pleasure when the child imitated successfully, so that there would be some reward for the child and some basis for instrumental learning, the first occurrence of the imitative response can be explained on the basis of simple association by contiguity. Let us then conclude that contiguity is a *sufficient* condition for the simple kind of imitation described above.

Where does all this lead us, as far as understanding role-taking is concerned? Are the concepts of imitation and role-taking synonymous? I think not. I would like to posit that all role-taking is imitation, but not all imitation is role-taking. Where does the distinction lie? *Role* behavior, as it is commonly defined, is behavior characteristic of a *position*. The role behavior associated with a given position is the behavior which others expect an incumbent of that position to display. Let me give an example of what could be considered clearly role behavior. Mrs. A came to visit Mrs. B, and Mrs. A brought with her her young baby whom Mrs. B had not seen before. During the visit, the baby lay on a blanket on the living room floor. After a time, it began to cry, and its mother, Mrs. A, picked it up and changed its diaper. A week later, a very similar scene was enacted in the same setting.

CHILD DEVELOPMENT

Mrs. B adopted a baby; Mrs. A, her good friend, rushed over to see it, this time leaving her own baby at home. The newly-adopted baby lay on a blanket on the living room floor, and after a time, began to cry. This time Mrs. B, not Mrs. A, picked up the baby, soothed it, and changed it. Why was Mrs. B's action different? The stimulus situation was the same—an unfamiliar baby crying, and another woman present who could have taken care of it. Her behavior was different because she now occupied a different *position*, the position of "mother," toward which she and others had very well-formed views about the behavior appropriate to the position. Her action then was *role* behavior because the behavior was not elicited except when she occupied the appropriate position. It will be seen that we need the concept of "role" in psychology for much the same reason that we need the concept "drive"—in both cases, the problem is that a given stimulus situation sometimes leads to a given response, and at other times does not. We employ the concept "drive" to designate the conditions under which a previously-learned response will be performed (e.g., the hungry animal will run to the food box when hungry and will not do so when satiated, although external conditions remain constant). At least one set of operations which make a difference in the response to a given stimulus situation are those of deprivation, and these are the antecedent operations which we relate to the concept "drive." The antecedent operations which can be related to *role* are not so clear. In the case of role behavior, the controlling condition which will determine whether previously-learned behavior is performed appears to be a categorical label, applied by others or by the person himself. This categorizing of oneself and others and the performance of the actions appropriate to one's category seems to work quite smoothly in many instances. For example, a man and a woman who are strangers to one another approach a revolving door; custom dictates that the woman shall step forward and the man back, and a collision is avoided when each person performs the action appropriate to his sex role.

Let us return to the problem of role-taking in children. A child's imitation of someone with whom he interacts may properly be called "taking the role of another" only if the action imitated is inappropriate for a child, and is appropriate instead for the occupants of some other position or status. Here is an example: A 2-year-old had a toilet accident. She looked sternly at her mother and said, "I'm 'prised at you!" in the tone of voice ordinarily used by her mother. Another 2-year-old in a similar predicament said to her uncle, who was taking care of her, "Naughty girl! You must come and tell mommy!" Here we have children making statements appropriate to the disciplinary function of a mother. The child is, then, at this moment, playing the mother role. We assume that much similar role-playing occurs covertly and overtly during solitary play when no adult is present to reinforce it. A possible explanation of these bits of role-taking behavior is, I think, that the stimulus situation makes the child think of what the mother would say in such a situation; that the auditory image of the

mother's phrase being said is sufficient on the basis of the circular reflex principle to get the child to say the phrase, either aloud or covertly. Please note an assumption implicit in this explanation. It is that, whenever we hear someone else saying something to us, *we learn it*, provided the necessary response units are already in our repertoire.² This means that the child will not be able to call to mind the mother's scolding phrase until he has mastered the words that make it up, just as we are not likely to remember and rehearse the phrases that are spoken to us in a foreign language. Inaccurate learning of a set of phrases may occur when the hearer does not possess the necessary response units. For example, a first-grade child being taught the salute to the flag learned, "One country, in the miserable, with liberty and justice for all," not possessing the response-unit "indivisible." We would expect that the features of a mother's behavior that a young child learns would be a quite oversimplified version of the original because of the limitations in the child's response repertoire. But, if a child does already have command of the response units, then he can rehearse the phrases that he hears and will be in a position to employ them when an appropriate situation occurs. I do not intend to imply that all phrases spoken in a child's presence are equally well learned, nor equally likely to be employed by him at a future date. I will try to specify later what some of the conditions may be that govern the degree of learning.

When a young child does imitate actions that are appropriate only for adults, he is likely to meet with a negative reaction from his parents. If a little girl tries to use her mother's lipstick, she will be told that she's only a little girl, and that lipstick is for grown women. A child who has been spanked may try to "spank" his parent on an occasion that seems to him to call for it, but this response is not likely to be maintained long, at least overtly, in view of the fact that most parents do not allow their children to be overtly aggressive toward adults. Thus, a child will soon learn that while certain items of grown-up behavior, such as tying one's own shoes, are appropriate for a child to perform as soon as he has mastered the manual skills involved, other items of behavior are not appropriate for a child. These are usually labeled according to the age and sex role to which they do belong, and the child is often explicitly taught that certain activities which are denied him now (such as driving a car) he will be allowed to perform when he is older. Let us use the term "adult role behavior" for the actions that are appropriate for the adults with whom the child interacts, but which the child may not, or cannot (because of physical limitations) perform in turn. Adult role behavior characteristically includes disciplinary actions directed toward a child who does something disapproved; controlling actions designed to protect a child from danger or protect other people and precious possessions from damage by the child; and a variety of nurturant

² The reader will note that a contiguity theory of learning is implied, although drives that can be aroused and reduced in fantasy (e.g., aggression) could presumably serve as the basis for reinforcement of covert responses.

CHILD DEVELOPMENT

actions, such as feeding the child, comforting him when hurt, putting on an extra blanket when he is cold, etc.³

During the early phases of a child's development, his own actions are most intimately bound up with the adult-role actions of his caretakers. Every one of the child's needs can be satisfied only through the mediating action of a caretaker. He cannot get his own food, walk where he wants to go, reach something he wants on a shelf, or dress himself; part of the sequence of actions he performs to satisfy any of these wants must be to elicit the necessary cooperative behavior from his caretaker; then he must orient himself to perform efficiently the reciprocal actions which the caretaker requires from him. For example, the following characteristic behavior was observed in a 2-year-old: He would always go to the right-hand side of his high chair when ready for a meal, holding himself in a characteristic position ready to be picked up. When he had been lifted into the chair, he would duck his head sideways, for the tray to be swung over his head; then he would lift his chin for the bib to be put under it, then bow his head down on the tray so the bib could be tied around the back of his neck. He was obviously prepared for a particular sequence of actions by his mother. He was utterly confused and upset when a new baby sitter tried to lift him up from the *left* side of the high chair, and tried to put his bib on before swinging the tray into place. The little boy and his mother had learned what might be described as two halves of the same habit. Until she was asked about it, the mother was quite unaware that she always performed her part of the habit in a stereotyped way; but actually the interaction of these two people was made smooth and effortless by virtue of the fact that the behavior of each was highly predictable by the other. I wish to postulate that, in the course of an often-repeated intimate interaction sequence like this one, each person not only is oriented toward what the other person will do but actually anticipates it to the extent of reproducing some elements of the other person's action covertly. Thus, the child who says, "I'm 'prised at you," when he has been naughty is anticipating the phrase his mother will say, and rehearsing it.

We are now able to specify what some of the conditions are which determine how thoroughly we incorporate into our own behavior repertoire the actions of other people with whom we interact. The first hypothesis is that the more frequently we interact with another person (and the more our own actions are interdependent with his) the more we must learn to anticipate what he will do if the interaction is to be harmonious; and, in anticipating him, we learn the content of his behavior. This hypothesis

³ The views enunciated here and in the following portions of this paper have grown out of continued discussion among the staff of the Laboratory of Human Development at Harvard. While the other staff members should not be held responsible for this particular formulation, the basic ideas expressed are a product of our joint work. The writer wishes to express particular indebtedness to John W. M. Whiting, Wesley Allin Smith, and Beatrice Whiting.

would predict that the worker and the first-line supervisor would learn each other's behavior more thoroughly than would two people who were *two* levels apart in a hierarchy. This rule applies to the learning that occurs when both members of the diad are present and serving as stimuli to one another. But we may also assume that a certain amount of spontaneous rehearsal of the other person's actions occurs when he is *not* present. Consider, for example, the case of a small child who falls down in the back yard and cuts his knee. While he runs to the house, crying for his mother, he thinks of his mother bandaging his knee, before she is actually there doing it. We are accustomed to dealing with expectations in S-R learning theory by labeling them as " r_G "—the anticipatory goal response. But note that " r_G " is a consummatory action—the child's own action; in the example just given " r_G " would consist of anticipatory elements of the process of *being bandaged*; in addition, I am now saying, there is an anticipatory response which includes at least some elements of the response of the other person; in this case, the child is practicing an anticipatory response of *bandaging*, as well as that of *being bandaged*.

I have said that the frequency with which we practice elements of another person's responses is a function of the frequency with which we interact with that person, and the intimacy with which our actions must be coordinated with his. But we suggest there is a more important condition governing the frequency of covert practice of another's responses, and this condition has to do with the power relationships between the two people concerned. Our second hypothesis is this: the more power another individual exercises over ego, the more ego will rehearse alter's actions in the absence of alter. John W. M. Whiting⁴ has used the phrase "control of resources" to describe more fully what kind of exercise of power will bring about covert practice of the role of another. A resource is something ego wants. According to Whiting's hypothesis, if the resource is mediated by alter, so that ego cannot satisfy his need except through a mediating action of alter's, then ego will more fully learn alter's characteristic behavior. Why should this be so? Presumably, because there will be numerous occasions upon which the arousal of a need in ego will provide the occasion for him to think about an action of alter's. Suppose a child is hungry; if he lives in a household where he is allowed to go to the refrigerator and get the ingredients for a sandwich, he will not engage in any covert role practice. If he is not allowed to do this, but must ask his mother, then, when he becomes hungry, he will think about asking his mother for a sandwich and will imagine a number of things she might say in response, "It's too soon before dinner," or, "If you've finished cleaning your room, you may have it." He may decide not to ask her for the sandwich after all, but meanwhile he will have practiced some of her characteristic responses, and added an increment to the habit strength of these responses in his own repertoire.

⁴ Unpublished working paper, Laboratory of Human Development, Harvard Univer.

CHILD DEVELOPMENT

If most things that an individual wants are not under his control but under the control of others, then presumably much of the vicarious trial-and-error that he engages in must involve his trying out various approaches to getting the help or avoiding the censure of others, and imagining the probable responses to these approaches. In our view the process of learning through role-playing the actions of others encompasses instances of "identification with the aggressor." It is difficult for Mowrer's autistic theory (4) to explain why a child repeats scolding statements to himself; they do not have secondary reinforcing power, but, on the contrary, have been paired with painful situations. Our position says that a child should covertly rehearse both the rewarding and the punishing actions characteristic of his parents, for both are highly relevant to him in guiding his plans about future actions.

Let us reiterate that we are *not* saying that the mind is like a tape recorder, automatically recording and retaining for future use all the phrases and actions that other people emit in ego's presence. We are saying that the learning process is highly selective, and that a law of frequency applies. That is, the probability of ego's using a particular phrase that he has heard someone else use may be very low or very high, depending in part upon the number of times he has rehearsed this phrase to himself. Furthermore, the conditions which will cause him to rehearse it are predictable.

The relations of status to role-taking should now be clear. Our first hypothesis was that whenever two people interact frequently, especially if the actions are mutually dependent in some way, both participants will learn something about the actions of the other. This would mean that the mother does some role-playing of her child's behavior, as well as vice versa. But the second hypothesis was that fantasy role-playing when the other person was not present was a function of power; according to this hypothesis, the worker should think about the behavior of his boss much more than the boss should imagine the behavior of his employees. And within the family, the young child must be almost wholly oriented toward the actions of his parents, since the parents mediate all resources for him, while the parents have much less reason to role-play the child in fantasy when he is not present. As a matter of fact, it is probable that learning by role-playing occurs more frequently in early childhood than at any other time of life, since the young child is almost completely dependent upon others; as the child grows older, he can fulfill more of his own needs without the mediation of another person, at which time his learning should be more of the direct instrumental kind.

Adult role behavior was defined earlier as behavior not deemed appropriate for children. Let us consider the fate of the adult-role actions learned by the child through fantasy role-playing. The child does not have an opportunity to practice these actions overtly. There may be a number of reasons why he cannot practice them. A major reason is that the appropriate stimulus situation does not arise. No one is directing child-like actions toward

him that would call for a reciprocal adult-like response. Secondly, a child cannot practice some adult-role actions overtly because of physical limitations (e.g., he is too small to lift another 2-year-old into a high-chair). And finally, if an opportunity did arise to perform an adult-role act and the child was physically capable of doing so, he might still be prevented by his adult caretaker on the grounds that the behavior is not appropriate for his age. So overt practice does not occur, but we assume that *covert* practice continues to occur so long as the child is in a dependent position in which the resources he needs are mediated by others. Now what about behavior that has been practiced covertly *only*, without opportunity for overt practice? Can such actions be said to have habit strength in the same way that conventional overt responses do? I think yes, but of course only those elements of a response that *can* be practiced covertly would acquire habit strength. There is little motor feedback from a covert response. For the acquisition of motor skills, feedback is of great importance, especially in the coordination of discrete motor elements. Therefore, it would appear that covert role-playing would be an inefficient way to learn motor skills. An example of this fact is seen in the case of an army inductee, who listens to his sergeant barking orders for close-order drill; the inductee mutters the orders to himself as the sergeant gives them, and can soon say them covertly in perfect cadence; but when he is required to lead the squad and actually deliver the orders in a loud voice he finds himself out of breath and is soon hopelessly out of rhythm. His covert practice never included coordinating his breathing with shouting the orders aloud. It would follow from this, for example, that a young child that has frequently had its diapers changed would acquire a tendency to try to change a crying baby, but would not learn how to fold a diaper, nor how to put one on so as not to be stuck with a pin in the process. But while covert practice may be inefficient for the acquisition of a motor skill, it should be perfectly adequate for the acquisition of *verbal* response tendencies.

Perhaps a more important limitation on the kind of learning that can occur via covert role-practice has to do with the perceptions of stimuli to which alter is responding. So far, I have implied that being able to reproduce some element of the responses characteristic of another person would constitute playing the role of that other person. I have also pointed out that very young children (under 2) can do this. Yet we have been told by Piaget and others that very young children are egocentric—that they can't put themselves in the place of someone else. Piaget has illustrated this fact in a number of ways. First there is his "brothers test," when he asks a child how many brothers he himself has, and then how many brothers one of his brothers has (5). A child can report accurately about how many brothers he himself has at a fairly early age, but it is not until middle childhood that most boys realize that they must count themselves as one of their sibling's brothers. As another example, Piaget recounts the difficulty young children

CHILD DEVELOPMENT

have imagining how an object would look to someone standing in a different place than the place where the child himself is standing. An example of this particular kind of egocentrism was observed in a young child who was quite capable of using parental words and phrases in play with dolls. This child was capable, then, in one sense, of playing the role of another person. Yet the following incident occurred: The child liked her milk to be quite cold; she took a sip from a glass of milk, and said, "Mommy, it's too warm." Her mother said, "No, it isn't dear, I just took it out of the refrigerator." The child held up the glass for her mother to see, and then ostentatiously took another sip of it, saying triumphantly to her mother, "See? It is warm!" She evidently believed that when she tasted something, her mother tasted it, too. The realization that other people are not experiencing the same stimuli that we are comes slowly; a young child, talking to someone on the telephone, will hold up an object for the person at the other end of the line to see, not realizing that the other person is in an entirely different visual environment.

We see, then, that while a young child may be able to reproduce the response of another person, this does not mean that he can put himself in another's place in the empathic sense. He may be making quite a different stimulus-response connection than his model is making. That is, he may connect his model's response (as represented in his own fantasy) to whatever set of stimuli are impinging upon himself at the moment, because he is not yet capable of discriminating his own cues from those stimulating others. Thus, for example, if a mother is cleaning out a closet and comes upon a garment with moth-holes and says, "Oh! How awful!" the small child playing in the room may practice this remark in relation to whatever he himself was looking at, and may learn nothing whatever about the evil qualities of moth-holes.

What does all this mean in relation to the kind of behavior that may be learned by covert role-playing? First, it means that a child may learn to respond to a different cue than that which elicited the mother's response; perhaps more importantly, he may *fail* to learn to respond to the cues that were actuating the mother's behavior. Let us give an example that illustrates why this failure may be important. Recently we were interviewing some children between the ages of 11 and 16 concerning their concepts of which kinds of behavior were adult-like and which child-like, and concerning their observations of children assuming adult roles. A high-school girl, when asked what she had observed about 12-year-olds acting as baby sitters, said, "Personally, I wouldn't trust them." When asked why, she said, "They're still in the 'soft' stage." Further exploration revealed that according to her observation, a very young baby sitter feels sentimental over little children, but would fail to respond to cues of danger. A young sitter could not tell, for example, when a 2-year-old was about to dash into the street, or about to fall, and consequently could not act the caretaking role effectively despite

the greatest feelings of solicitude for the child. I wish to suggest that this situation results from the young sitter's having learned only certain aspects of her mother's behavior in earlier years when she was herself a child receiving maternal care. She could hear, and practice, her mother's statements toward herself as a child; she could even practice her mother's protective actions; but she did not necessarily notice the cues that set off these actions in her mother. She had no practice, for example, in seeing that a particular facial expression in a child presaged a dangerous action that would call for an instant maternal response. A child has no training in responding to this kind of cue because he cannot see his own face. Instead, he *does* have training in the cue elements that are internal. Suppose, for example, that a child has certain internal feelings when he is about to mark on a wall. He can feel his own arm reaching out for the pencil, etc. This impulse will serve as a cue to remind him of the things his mother will probably say if he does it—he hears, in his mind's ear, the scolding phrase, "You naughty boy, give me that pencil this instant!," thus practicing this phrase and increasing the probability that he will one day say it himself in an appropriate situation. But, in the future, what cue will cause him to say it? We know that organisms do learn to respond to cues to which a model has responded. Church (1), in an experiment with rats, trained follower rats to make the same turns in a maze as the leader rats made. The leaders were following previously learned visual cues. After a number of runs, the follower rats were put in the maze alone, and were able to make some use of the visual cues the leaders had previously been following. In a similar vein, in studying the material that movie-viewers learned from a movie, Maccoby and Wilson (2) found that if there were two leading characters, and the viewer chose one of them as the person who became for him the principal protagonist of the story, he remembered later somewhat more of the cues relevant for his chosen hero's actions than those relevant for the actions of the other leading character. But we are suggesting that this learning will occur primarily with respect to those cues which stimulate both the observer and the person he identifies with in the same way and at the same time.

I am also suggesting that a child will learn to produce the adult-role response more dependably to *internal* cues than to external ones. Thus, in the case of the 12-year-old baby-sitter, we would expect that if the sitter *herself* were tempted to run into the street, or felt *herself* endangered by an oncoming car, then she would make a protective gesture toward her young charge. We see this clearly in the case of a very young child; if he says to his doll, "You naughty doll, you wet your pants," the experienced mother will urge the child himself to go to the bathroom. She rightly assumes that the child's disciplinary remark toward the doll was elicited by the cues of the child's own impending deviation.

This brings up a central point, having to do with the development of attitudes toward the self. If a child learns the adult-role behavior charac-

CHILD DEVELOPMENT

teristic of his parents, he may be expected to manifest the behavior in two sets of circumstances: when another person performs child-like actions toward him (e.g., when another child is hurt and needs solace, or when another child breaks a rule and needs discipline), or when he himself performs these child-like actions or has the impulse to do so. Suppose a child has parents who characteristically react by withdrawing their love from him when he does something they disapprove of. If he learns their characteristic mode of response to the deviation of another person, we would expect him to react to another child who breaks a rule by refusing to play with the other child. But, in addition, when he himself deviates, he will respond to his own impulse by withdrawing love from himself—a phenomenon we more commonly refer to as a loss of self-esteem. Thus, the parents' attitudes toward the child, and the parents' techniques of dealing with the child, will be reflected in the child's self-attitudes as well as in his attitudes and behavior toward others. This idea will not be a new one, to followers of psychoanalytic theory, nor to those of G. H. Mead. But it does imply specifically, for example, that we should not regard the trait of extrapunitive-ness and that of intropunitive-ness as being opposite ends of a scale. Rather, if a child has punitive parents, we would expect that, if the conditions are such that the child has engaged in a good deal of covert practice of the parents' actions, he will in the future be punitive toward others who deviate, and will be likewise punitive toward himself if he transgresses.

We must guard against having a theory that predicts more homogeneity between generations than actually exists. So far, I may have implied that most children will behave just as their parents do. Let us point out some specific reservations about such a generalization. First of all, parents are not the only people who control resources the child wants. There are usually a succession of important figures who perform this function. First, the mother assumes primary importance with respect to the issues of early childhood; feeding, toilet training, and dependency. The father may play a more important role with respect to the control of aggression, especially for sons, and later controls one of the most important resources of all: information about skills the boy wants to acquire. Teachers play an important part, as do older brothers and sisters, and finally of course, the individual's spouse. We would expect that the growing individual would add behavior items to his repertoire from each of these people with whom he interacts. It would also follow that for most people the behavior they show when a young child asks them for food will be like that of their mothers, while their behavior in response to other kinds of instigations may have been learned through interaction with different agents. Thus, we are not talking about "identification" with any *single* individual in the child's life history, and no child should be a mirror image of any single adult.

Another factor which limits the similarity between a child's actions and those of his parents is instrumental learning. The action tendencies learned

ELEANOR E. MACCOBY

through covert role-playing may be in conflict with the action tendencies learned through the reward and punishment of overt actions. Consider, for example, the case of a boy who lives in a neighborhood where there is only one other child his age. The other child pushes him around mercilessly until one day he chances to push back hard, and the other child stops harassing him. He will learn aggression as a successful instrumental response to the attack of another person. Suppose his mother is the kind of person who responds to attack by withdrawal; the child will have learned this tendency, too, through covertly practicing her behavior. It is difficult to predict what will be the outcome of these conflicting action tendencies; perhaps, the child will be more aggressive than his mother, but less so than the success of his overt aggressive behavior would lead one to expect; or perhaps the child will begin to discriminate a set of situations where aggression is appropriate from a set where withdrawal is appropriate. In any case, he will not be precisely like his mother with respect to this area of behavior. As a matter of fact, we might expect that there would often be some conflict involved in the overt expression of behavior learned through role-playing the parent of the opposite sex. Particularly for boys, the rewards for masculine behavior and the disapproval of feminine behavior are such that the strictly feminine items of behavior that he has learned through role-playing his mother will never become manifest, but will be latent all his life.

Some children have much more opportunity than others to practice overtly, as they are growing up, the adult-role actions they have learned through interaction with their parents. Some children, for example, are allowed to discipline their younger siblings, others are not. If a child does have opportunity for overt practice of adult-role actions, we should expect that the tendencies learned through covert role-playing of the parents would be either strengthened or weakened, depending upon the response of others to the child's adult-like behavior. If he tries to boss other children and they ostracise him for it, his tendency to imitate a dominant parent would be weakened. We hypothesize that a child's adult-like action will be most similar to that of his parents on the first occasion when the action becomes overt. After this, the events which follow the response will often cause divergence of the child's response tendencies from that of his parents.

In summary, we have proposed that a child acquires a repertoire of actions by practicing covertly the actions characteristic of the adults with whom he interacts most frequently and who control the resources that he needs. Certain of the response-tendencies thus acquired may not manifest themselves overtly until a much later time. Specifically, actions which are part of adult-role behavior (e.g., to discipline a child) will remain latent until a situation arises in which the individual can appropriately play the adult role, sometimes even until the child himself becomes a parent. We have suggested that not all features of parental behavior are equally well learned by the child through covert role-practice; he should learn verbal

CHILD DEVELOPMENT

behavior more efficiently than motor skills by this means, and may learn the responses of others while failing to learn the cues which guided the responses. And finally, we have suggested that covert role-playing is a means of learning not only adult-like social actions directed toward others, but of learning reactions toward the self.

REFERENCES

1. CHURCH, R. M. Transmission of learned behavior between rats. *J. abnorm. soc. Psychol.*, 1957, 54, 163-165.
2. MACCOBY, ELEANOR E., & WILSON, W. C. Identification and observational learning from films. *J. abnorm. soc. Psychol.*, 1957, 55, 76-87.
3. MILLER, N. E., & DOLLARD, J. *Social learning and imitation*. New Haven: Yale Univer. Press, 1941.
4. MOWRER, O. H. On the psychology of talking birds. In *Learning theory and personality dynamics*. New York: Ronald Press, 1950. Pp. 689-707.
5. PIAGET, J. *Judgment and reasoning in the child*. New York: Harcourt, Brace, 1928.
6. PIAGET, J. *Play, dreams, and imitation in childhood*. New York: Norton, 1951.

STUDIES ON ANXIETY: I. ANXIETY AND PERFORMANCE ON PSYCHOMETRIC TESTS VARYING IN COMPLEXITY¹

BEEMAN N. PHILLIPS,* F. J. KING, and CARSON MCGUIRE

The University of Texas

There have been a number of studies of the performance and personal characteristics of anxious and nonanxious Ss. Comparisons of these groups have been made in conditioned learning experiments (3, 13, 34, 35, 36, 37, 38, 39), and a number of studies of complex learning also have been reported (4, 8, 16, 20, 21, 24, 25, 26, 27, 28, 28, 41). Other studies have investigated the personality correlates of anxiety (2, 5, 10, 11, 15, 22, 23, 30, 31, 32) and the validity and comparability of various measures of anxiety (1, 6, 7, 14, 18, 19, 33, 40).

Most of these studies are based on Hull's theoretical formulation of response strength and drive (17). Hull theorized that sE_R (response evocation potential) is a multiplicative function of sH_R (habit strength) and D (drive), or that all habit tendencies aroused by a stimulus situation are multiplied by the total drive level. Using this formulation, Spence and his associates developed several expectations with respect to the performance of anxious and nonanxious subjects in conditioning and complex learning experiments (33). One deduction was that, if the learning situation involved a single or dominant response, anxious Ss would show superior performance because this response tendency would be enhanced by the higher drive level of anxious Ss. The other deduction was that in a complex learning situation a number of response tendencies would be aroused, and in this situation a higher drive level would not necessarily aid performance.

* Department of Educational Psychology, University of Texas, Austin 12.

¹ The authors are collaborators on the Human Talent Research Project at the University of Texas, which is supported by a grant from the U. S. Office of Education (Project No. 025-6431) and is directed by Dr. Carson McGuire. The data reported in this study were collected as part of the project.

CHILD DEVELOPMENT

If the correct response is dominant over other competing responses, performance would be facilitated, but if the correct response is not dominant, a high drive level would have the effect of impeding performance.

Because most studies of anxiety have used adults or college students as *Ss*, there has been a lack of information on the effects of anxiety on the performance of children. However, McCandless and his associates recently developed an anxiety scale for children and duplicated some of the learning studies done with adults (4, 22, 23, 27). But there also is a lack of information concerning the effect of anxiety on psychometric test performance (28). In addition, the results of Grice (11) and McCandless and Castaneda (23) raise the question of whether differences in intelligence account for differences in the performance of anxious and nonanxious *Ss* on complex learning tasks.

In this study anxious and nonanxious adolescent *Ss* who differed in intelligence were used, and their performance on selected psychometric tests was compared. In terms of the theoretical formulation, a different order of relationship between anxiety and performance would be predicted for the various tests included in the battery, since judgmental and empirical evidence reveals that these tests differ in degree of complexity (number of competing responses). On the basis of the rationale, it is hypothesized that the magnitude of the computed partial correlation coefficients (partialing out the possible effects of intelligence) will vary inversely with the level of complexity. The relationship will be highest for the simpler and easier tests and progressively lower for the more complex tests.

PROCEDURE

The subjects for this study were chosen from approximately 1400 seventh grade students in four Texas communities who make up the population for the Human Talent Research Project. Nine hundred and fifty-six students had taken all of the tests used in the study, and they comprise the sample from which the correlations were obtained.

Of the tests used, six have been described elsewhere. These are the McCandless Anxiety Scale (23), Mutilated Words, Gestalt Completion, Four Letter Words, Writing Lack (9), and the California Test of Mental Maturity. The remaining tests are somewhat less well-known and are briefly described here:

a. Gestalt Transformation—This test includes 24 items, each consisting of one problem and five possible answers. In each case, the possible answers are names of common objects, and the subject is asked to indicate from which one he could take a *part* in order to solve the problem.

b. Vocabulary Completion—For each of 80 items a short definition is given and the subject is asked to supply a word which fits it and which begins with a given letter. Both this test and the Gestalt Transformation test were taken from Guilford's Creativity battery (12).

c. Dotting Test—The subject is asked to put three pencil dots in each one in a series of circles as quickly as possible. Speed is of more importance than accuracy. This and the remaining three tests were developed for use in the U. S. Air Force.²

d. Writing X's—Here the subject is presented with a series of squares, each of which has smaller squares in its four corners. The instructions are to put an X in each small square so that no part of it is outside the boundaries.

e. Discrimination Reaction Time—Each item on this test consists of a square containing one black circle and one white circle. To the right of each square are four spaces arranged to indicate the positions "above," "below," "left," and "right." The subject is told to put a check mark in the appropriate space according to the position of the white circle in relation to that of the black one.

f. Copying—In this test the subject is asked to copy each of 36 figures. Each figure is a line containing three angles; there is variation both in length of line between angles and in acuteness of angles. For each example, figures must be copied within a square grid of 25 dots so that all lines begin and end on dots and all angle "points" coincide with dots.

The diversity of the population and numerous other conditions in the four communities made it desirable to standardize as much as possible the circumstances directly relating to the administration of the tests. For this reason, the entire testing program was carried out by a team of six experienced psychometrists.

To obtain a rank ordering of the tests on the basis of complexity, four members of the testing team and one faculty sponsor of the project were asked to rank the 10 tests with regard to the number of competing responses aroused in answering the test items. The degree of agreement for the five sets of ranks was determined by the coefficient of concordance; a W coefficient of .86 indicated significant agreement beyond the 1 per cent level of confidence.

RESULTS AND DISCUSSION

In this study the proposition that anxiety differentially affects performance on psychometric tests purportedly varying in complexity was investigated. It was hypothesized that anxiety would be more highly related to performance on simpler, less complex tasks where there are fewer competing responses than on complex tasks with many competing responses.

The data used to test the hypothesis are shown in Table 1 which shows the correlation coefficients between anxiety and performance on psychometric tests ranked according to complexity. Partial correlation coefficients in which

² Used by special permission of the Air Research & Development Command, United States Air Force, Lackland Air Force Base, Texas.

CHILD DEVELOPMENT

TABLE I

CORRELATION COEFFICIENTS BETWEEN ANXIETY, INTELLIGENCE, AND
SCORES ON 10 PSYCHOMETRIC TESTS ($N = 956$)

Test	Rank Order (Complexity)	r_{12}	r_{13}	$r_{12.3}$
Dotting	1	-.087	.263	-.039
Writing Lack	2.5	-.021	.130	.004
Marking X's	2.5	.025	.045	.034
Discrimination Reaction Time ..	4	-.138	.344	.079*
Short Words	5	-.048	.339	.030
Copying	6	-.135	.501	-.047
Vocabulary Completion	7	-.187	.698	-.078*
Gestalt Transformation	8	-.133	.476	-.050
Gestalt Completion	9.5	-.127	.388	-.059
Mutilated Words	9.5	-.052	.418	.030

NOTE.—1 = score on test; 2 = anxiety score; 3 = intelligence score.

* Significant at .05 level.

the effects of intelligence on test performance have been eliminated are also shown.

In examining Table 1 it is evident that the data do not support the hypothesis. Only two coefficients reach significance at the 5 per cent level, and these are statistically significant only because of the very large size of the sample.

In view of the number of studies recently reported in the psychological literature which have found results supporting the hypothesis, some speculations concerning possible reasons for the deviant findings of this investigation appear in order.

One possible explanation is that the relationship between anxiety and performance is curvilinear. For example, high anxiety may impede performance on complex tasks because there is too much drive, while low anxiety may impede performance because there is too little drive.

Sex is another factor which may need consideration. If adolescent males and females differ in average anxiety level, or if anxiety has differential effects on their performance, the failure to study the sex variable might lead to a suppression of anxiety and performance relationships. Another personal characteristic of Ss which may need to be considered is social class status. In view of the large number of anxiety studies in which college students have been used as Ss, it is possible that this factor accounts for some of the difference between the results of this and other studies.

The conditions under which the tasks were administered may also be a significant factor in this investigation. In many past studies Ss have been tested individually, whereas in the present study Ss were tested in large

groups. If an individual testing situation is more ego-involving than a group testing situation, this raises a question of whether an anxious S's performance is impeded more by an individual, personal situation than by a group, impersonal testing situation.

In summary, no evidence was found to support the hypothesis that anxiety differentially effects performance on tasks differing in difficulty and complexity. In addition, possible reasons for these "negative" results were explored and several ideas were advanced. These ideas are being subjected to further research.

REFERENCES

1. BECHTOLDT, H. P. Response defined anxiety and MMPI variables. *Proc. Iowa Acad. Sci.*, 1953, 60, 495-499.
2. BENDIG, A. W. Age, sex, and the manifest anxiety test. *J. consult. Psychol.*, 1954, 18, 16.
3. BITTERMAN, M. B., & HOLTZMAN, W. H. Conditioning and extinction of the galvanic skin response as a function of anxiety. *J. abnorm. soc. Psychol.*, 1952, 47, 615-623.
4. CASTANEDA, ALFRED, et al. Complex learning and performance as a function of anxiety in children and task difficulty. *Child Development*, 1956, 27, 327-332.
5. ERIKSEN, C. W. Some personality correlates of stimulus generalization under stress. *J. abnorm. soc. Psychol.*, 1954, 49, 561-565.
6. ERIKSEN, C. W., & DAVIDS, A. The meaning and clinical validity of the Taylor anxiety scale and the hysteria-psychasthenia scales from MMPI. *J. abnorm. soc. Psychol.*, 1955, 50, 135-137.
7. FARBER, I. E. The role of motivation in verbal learning and performance. *Psychol. Bull.*, 1955, 52, 311-327.
8. FARBER, I. E., & SPENCE, K. W. Complex learning and conditioning as a function of anxiety. *J. exp. Psychol.*, 1953, 45, 120-125.
9. FRENCH, J. W. (Ed.) *Manual for kit of selected tests for reference aptitude and achievement factors*. Princeton, N.J.: Educational Testing Service, 1954.
10. GOODSTEIN, L. D. Interrelationships among several measures of anxiety and hostility. *J. consult. Psychol.*, 1954, 18, 35-39.
11. GRICE, G. R. Discrimination reaction time as a function of anxiety and intelligence. *J. abnorm. soc. Psychol.*, 1955, 50, 71-74.
12. GUILFORD, J. P. Relation of intellectual factors to creative thinking in science. In C. W. Taylor (Ed.), *1955 University of Utah research conference on the identification of creative scientific talent*. Univer. of Utah Press, 1956, Pp. 69-95.

CHILD DEVELOPMENT

13. HILGARD, E. R., JONES, L. W., & KAPLAN, S. J. Conditioned discrimination as related to anxiety. *J. exp. Psychol.*, 1951, 42, 94-100.
14. HOLTZMAN, W. H., CALVIN, A. D., & BITTERMAN, M. E. New evidence for the validity of Taylor's manifest anxiety scale. *J. abnorm. soc. Psychol.*, 1952, 47, 852-854.
15. HOLTZMAN, W. H., ISCOE, I., & CALVIN, A. D. Rorschach color responses and manifest anxiety in college women. *J. consult. Psychol.*, 1954, 18, 317-324.
16. HUGHES, J. B., SPRAGUE, J. L., & BENDIG, A. W. Anxiety level, response alteration, and performance in serial learning. *J. Psychol.*, 1954, 38, 421-426.
17. HULL, CLARK L. *A behavior system*. New Haven: Yale Univer. Press, 1952.
18. KERRICK, JEAN S. Some correlates of the Taylor manifest anxiety scale. *J. abnorm. soc. Psychol.*, 1955, 50, 75-77.
19. JESSOR, RICHARD, & HAMMOND, KENNETH. Construct validity and the Taylor anxiety scale. *Psychol. Bull.*, 1957, 54, 161-170.
20. LAZARUS, R. A., DEESE, J., & HAMILTON, R. Anxiety and stress in learning: the role of intraserial duplication. *J. exp. Psychol.*, 1954, 47, 111-114.
21. LUCAS, J. D. The interactive effects of anxiety, failure and intraserial duplication. *Amer. J. Psychol.*, 1952, 65, 59-66.
22. McCANDLESS, BOYD R., *et al.* Anxiety in children and social status. *Child Developm.*, 1956, 27, 385-391.
23. McCANDLESS, BOYD R., & CASTANEDA, ALFRED. Anxiety in children, school achievement, and intelligence. *Child Developm.*, 1956, 27, 379-382.
24. MALTZMAN, I., FOX, J., & MORRISETT, L. Some effects of manifest anxiety on mental set. *J. exp. Psychol.*, 1953, 46, 50-54.
25. MATARAZZO, J. D., ULETT, G. A., & SASLOW, G. Human maze learning performance as a function of increasing levels of anxiety. *J. gen. Psychol.*, in press.
26. MONTAGUE, B. K. The role of anxiety in serial rote learning. *J. exp. Psychol.*, 1953, 45, 91-96.
27. PALERMO, DAVID S., *et al.* The relationship of anxiety in children to performance in a complex learning task. *Child Developm.*, 1956, 27, 333-337.
28. PICKREL, EVAN W. The differential effect of manifest anxiety on test performance. *J. educ. Psychol.*, 1958, 49, 43-46.
29. RAMOND, C. Anxiety and task as determiners of verbal performance. *J. exp. Psychol.*, 1953, 46, 120-124.

30. SMOCK, C. D. The influence of psychological stress on the "intolerance of ambiguity." *J. abnorm. soc. Psychol.*, 1955, 50, 177-182.
31. SMOCK, C. D. The influence of stress on the perception of ambiguity. *J. abnorm. soc. Psychol.*, 1955, 50, 354-356.
32. SMOCK, C. D. The relationship between "intolerance of ambiguity," generalization, and speed of perceptual closure. *Child Developm.*, 1957, 28, 27-36.
33. SPENCE, K. W. *Symposium on relationships among learning theory, personality theory and clinical research*. New York: Wiley, 1953.
34. SPENCE, K. W., & BEECROFT, R. S. Differential conditioning and level of anxiety. *J. exp. Psychol.*, 1954, 48, 399-403.
35. SPENCE, K. W., & FARBER, I. E. The relation of anxiety to differential eyelid conditioning. *J. exp. Psychol.*, 1954, 46, 127-134.
36. SPENCE, K. W., FARBER, I. E., & TAYLOR, BLAINE. The relation of electric shock and anxiety to level of performance in eyelid conditioning. *J. exp. Psychol.*, 1954, 48, 404-408.
37. SPENCE, K. W., & TAYLOR, JANET A. Anxiety and strength of UCS as determiners of the amount of eyelid conditioning. *J. exp. Psychol.*, 1951, 42, 183-189.
38. SPENCE, K. W., & TAYLOR, JANET A. The relation of conditioned response strength to anxiety in normal, neurotic, and psychotic subjects. *J. exp. Psychol.*, 1953, 45, 265-272.
39. TAYLOR, JANET A. The relationship of anxiety to the conditioned eyelid response. *J. exp. Psychol.*, 1951, 41, 81-92.
40. TAYLOR, JANET A. A personality scale of manifest anxiety. *J. abnorm. soc. Psychol.*, 1953, 48, 285-290.
41. TAYLOR, JANET A., & SPENCE, K. W. The relationship of anxiety level to performance in serial learning. *J. exp. Psychol.*, 1952, 44, 61-64.

The history of the United States is a story of the growth of a nation from a collection of small colonies to a great republic. It is a story of the struggles of the people for freedom and justice, and of the triumphs of the American spirit. The story begins with the first settlers, who came to the New World in search of a better life. They found a land of opportunity, but also a land of hardship. They fought for their rights, and they won. They built a nation that was free and independent, and that was the envy of the world. The story continues with the growth of the nation, and the struggles of the people for a better life. It is a story of the American dream, and of the American way of life. The story ends with the present, and the future of the United States. It is a story of hope, and of the promise of a better future for all.

EFFECTS OF MOTIVATING INSTRUCTIONS ON REACTION TIME IN GRADE SCHOOL CHILDREN¹

WAYNE A. OWEN*

State University of Iowa

There is a striking paucity of data concerning the main and interactive effects of motivational factors on reaction time (RT) in both children and adults. The most widely cited study of this type is one by Johanson (5), who found, with three adult Ss, that simple RT was accelerated by punishing Ss with electric shock for relatively slow reactions or by simply furnishing knowledge of results after each trial. Rosenbaum (6), studying choice RT, found that threat of strong shock for slow reactions led to an improvement in the performance of college students. Threat of weak shock was not significantly effective. However, Farber and Spence (4), also working with college students, found no significant relations between threat of strong shock and movement time required for simple and choice reactions. These authors suggest that the effects of stress may interact with other variables which are as yet unknown.

In a recent study, Blackburn (2) has investigated the effects of three kinds of motivating instructions (standard, urging, relaxing) on choice RT in nonpsychotic patients with supratentorial cerebral disease and a comparable group of patients with diverse neurological and medical complaints but with no cerebral disease. Initially, 30 visual choice RT trials were administered to each S under standard conditions. This was followed by differential instructions and by 30 postinstructional trials identical with the

* Box 351, Connecticut State Hospital, Middletown.

¹ Adapted from an M.A. thesis, done under the direction of Prof. Arthur L. Benton. The writer is greatly indebted to Prof. Benton for his guidance of the study and to Profs. A. Castaneda, I. E. Farber, and D. W. Norton for valuable advice and criticism. He also wishes to thank Mr. B. W. Garner and Mr. D. Seavy of the Iowa City school system for their cooperation in making available subjects for the study.

This investigation was supported by a research grant (B-616) from the National Institute of Neurological Diseases and Blindness, of the National Institutes of Health, U. S. Public Health Service.

CHILD DEVELOPMENT

initial 30 trials. The criterion score for each *S* was the difference in total time scores between his 30 preinstructional and 30 postinstructional trials. After standard instructions, the control *Ss* showed a slight acceleration in RT while the brain-damaged *Ss* showed a slight retardation, this difference in the trends of the two groups being significant. Under the urging retest instructions, both groups showed significant acceleration in performance as compared with both their initial performances and their performances under the other two instructional conditions. The acceleration was somewhat greater in the brain-damaged than in the control group, the difference in gain scores of the two groups approaching significance. Under the relaxation instructions, both groups showed a slight acceleration in performance. In the control group this acceleration was identical in amount with that shown under the standard retest instructions. In the brain-damaged group, however, this acceleration contrasted with the slight retardation shown under the standard retest instructions, the difference in trends approaching significance.

The present study represents an extension of this type of investigation to the performances of normal children. Using essentially the same procedure as did Blackburn, the effects of the three types of motivating instructions on choice RT in grade school children were examined. A secondary purpose was to examine the effects of the sex factor on the choice RT performance of the children.

METHOD

Subjects

Sixty grade school children were given two consecutive administrations of a test involving a choice reaction to one of two lights. The first administration was always given under standard instructions. The second administration was given under one of three sets of motivating instructions (standard, urging, relaxing).

The sample consisted of 30 boys and 30 girls between the ages of 7 years, 7 months and 9 years, 4 months. The mean age of the boys was 8 years, 6 months and that of the girls was 8 years, 7 months. Within the limits imposed by the maintenance of equality with respect to sex, one third of the children were randomly assigned to each instructional group. Thus, there were 20 *Ss* (10 boys, 10 girls) in each instructional group.

Apparatus

The reaction time apparatus employed was essentially identical with that used in previous investigations by Benton and Blackburn (1, 2, 3). It provided for the presentation of a warning buzzer to the *S* whenever a switch on *E*'s side of the apparatus was thrown to the "ON" position. The buzzer sounded for 2 seconds. Simultaneously with its cessation, one of two 6-watt light bulbs mounted on an upright metal panel, 10 inches apart, 7 inches

above the table top, and 6 inches in front of *S*, was lighted and a standard electrically driven clock was started. Two microswitches were mounted directly below the lights, 7 inches apart and $3\frac{1}{2}$ inches from the table's edge, on a horizontal metal plate which was rigidly attached to the vertical panel and extended to the edge of the table. Depression of the microswitch on the same side as the lighted stimulus extinguished the light and stopped the electric clock, and its reading in .01 second units (giving the elapsed time between the onset of light and depression of switch) was recorded manually. The clock was then manually reset to zero time. Depression of the incorrect microswitch had no effect on the clock. It served only to close the circuit to a small neon light on *E*'s side, thus indicating to him any attempt to press both switches indiscriminately. A second switch on *E*'s side of the apparatus determined which light (either right or left) would go on and which microswitch would stop the electric clock. The apparatus was reset manually by *E* after each trial.

Administration

S was seated comfortably in front of the light panel of the apparatus, with his arms resting on the table. *E* sat across the table on the other, or control, side of the apparatus. The following instructions were given:

This is a game to see how fast you can press down the key when you have to make a choice between the two lights. This time, you place the forefinger of your right hand on this key (point) and the forefinger of your left hand on this key (point) and rest them there without pressing down. I'll start a buzzer from over here again; and, when the buzzer stops, one of the two lights on the panel in front of you will flash on. Keep watching for the light to flash, and, when you are sure the right light is on, press down the right key (point) and *not* the left as quickly as you can. When you are sure the left light is on, press down the left key (point), but not the right, again as quickly as you can. Never press down both keys at one time. It's either one or the other. Is that clear? (Repeat relevant portions as necessary.) Now we'll try some practice (5 practice trials). That's the end of the practice; now we'll start the regular game (30 test trials).

Any response completed in .100 second or less or any depression of the key on the incorrect side, regardless of whether the correct key was depressed before or after, was considered to be an anticipatory error and an additional trial was given.

Thirty simple RT trials always preceded the choice RT procedure. In the latter, *S* was instructed to rest the index fingers of his right and left hands on the right and left microswitches and to press the appropriate switch when the right or left light flashed. The light stimulus was preceded by the warning buzz of 2 seconds duration. Lights on *E*'s side of the apparatus indicated the key which was pressed so that double pressing could be noted. Five practice trials were followed by 30 experimental trials, consisting of 15 stimulations on the right and 15 stimulations on the left side presented in a predetermined random order.

CHILD DEVELOPMENT

Following a rest period of approximately 2 minutes, one of the three following sets of instructions was given to precede a retest administration consisting of 3 warm-up and 30 experimental trials (15 presentations of each light in the reverse order of the initial administration presentation):

a. *Standard or task-orienting instructions:* Now let's try this game again. We'll do the same thing that we did before. Let's have some practice first.

b. *Urging and evaluating instructions:* Now let's try this game again. We'll do the same thing we did before, only this time I want you to see just how very fast you can press down the key. I've just figured how you did last time, and you should be able to do better than that. You should be able to press down the key faster than that. This time, after you press down the key each time, I will tell you whether you did better or worse than the first time. Be sure and pay close attention so that you can keep doing better each time. Knowing how you are doing each time should help you do better. Let's have some practice first. (After each trial, *E* reported whether *S* was doing better or worse than the average of his first set of trials. For each trial *S* was urged to do even better.)

c. *Relaxing and reassuring instructions:* Now let's try this game again. We'll do the same thing we did before, only this time I think you would do better if you just take it easy and relax. Remember, this is just a game. You've been doing fine so far, but just relaxing seems to help everybody. Still try to press down the key as quickly as you can, but don't worry about how you are doing. Relax, take it easy as much as possible while you are waiting for the light to come on. Let's have some practice first. (After practice, *S* is assured that he is doing fine, even better than before, and that he should continue to relax. Reassurance is repeated after trials 5 and 20 or whenever *S* makes spontaneous inquiry.)

RESULTS

Analysis of Initial Scores

A preliminary analysis of initial test scores (mean of the time measures for the first 30 trials under standard conditions) was performed in order to evaluate any differences in absolute level of performance before instructions treatments were imposed. Mean initial choice RT for the boys was 460 msec. for the standard instructions group, 479 msec. for the urging instructions group, 465 msec. for the relaxing instructions group, and 468 msec. for the combined instructional groups. Corresponding initial scores for the girls were 460 msec., 483 msec., 463 msec., and 469 msec. For the total group the mean initial scores were 460 msec., 481 msec., and 464 msec., respectively. A 2×3 factorial analysis of variance indicated that there were no significant differences between boys and girls ($F < 1.00$) and that there were no significant differences between instructional groups ($F < 1.00$) with respect to initial RTs.

Analysis of Criterion Scores

The criterion measure utilized in the analysis of the gain scores was the mean of the 30 time measures *before* instructional treatments minus the mean of the 30 time measures *after* instructional treatments. Criterion score

WAYNE A. OWEN

TABLE I

MEAN CRITERION SCORES AND SD (in msec.) FOR EACH SEX-INSTRUCTIONAL SUBGROUP ($N = 10$) AND EACH INSTRUCTIONAL GROUP ($N = 20$)

Instructions Group	BOYS		GIRLS		TOTAL	
	M	SD	M	SD	M	SD
Standard	-11	28	-20	27	-16	28
Urging	83	53	85	30	84	43
Relaxing	2	18	-1	41	1	32

means and their standard deviations for each of the six sex-instructional subgroups and the three instructional groups are presented in Table 1.

In inspecting this table, it should be kept in mind that a positive value indicates a *decrease* and a negative value an *increase* in RT. It will be noted that for each sex group and for the total group there was an apparent retardation in RT with repetition under standard conditions and an apparent acceleration under urging instructions while RT remained the same under relaxing instructions. These impressions were confirmed by statistical analysis of the findings. A 2×3 factorial analysis of variance indicated a significant instructional effect ($F = 18.0, p < .001$), a nonsignificant sex effect ($F < 1.00$), and a nonsignificant sex-instructions interaction ($F < 1.00$). The postinstructional RT of the standard instructions group was significantly slower than the preinstructional RT, as estimated by a t test for related samples ($t = 2.3, 19 \text{ df}, p < .05$). Similarly, the postinstruction RT of the urging instructions group was significantly faster than the preinstructional RT ($t = 8.5, 19 \text{ df}, p < .01$). On the other hand, the difference in RT under pre- and postinstructional conditions in the relaxing instructions group was nonsignificant ($t = .07, 19 \text{ df}$).

A comparison of the three instructional conditions indicated that the effect of urging instructions was significantly different from those of either the standard instructions ($t = 8.19, 54 \text{ df}, p < .01$) or the relaxing instructions ($t = 6.91, 54 \text{ df}, p < .01$). However, the difference in the effects of the standard and relaxing instructions was nonsignificant ($t = 1.27, 54 \text{ df}$).

DISCUSSION

The findings show that specifically defined motivating instructions have differential effects on RT in school children. Retest under standard instructions results in a significant retardation in RT. In contrast, urging instructions produce a significant acceleration in RT. The effect of special relaxation instructions is essentially to maintain speed of reaction at its preinstructional level. The accelerating effects of the urging instructions are demonstrably different from those of either the standard or relaxation in-

CHILD DEVELOPMENT

structions. However, the standard and the special relaxation instructions are not demonstrably different in their effects.

When these observations are compared with those of Blackburn on adult patients, it is seen that there is a high degree of consistency in the results with, however, one rather interesting difference. Blackburn's control patients showed a slight and nonsignificant acceleration in RT following standard retest instructions while his brain-damaged patients showed a slight and nonsignificant retardation in RT following the same instructions, the difference in the trends of the two groups being significant. It is evident that the children in this study resemble Blackburn's brain-damaged patients in that they too showed a retardation, rather than an acceleration, in RT following standard retest instructions. In both studies, a significant accelerative effect of special urging instructions is shown. The special relaxation instructions also have the same apparent effect, i.e., a slight and nonsignificant accelerative effect which contrasts with the retarding effect of standard retest instructions in the children and in Blackburn's brain-damaged patients.

Thus, an over-all comparison with Blackburn's findings indicates that normal children differ from adults in that the children show a negative practice effect under standard instructional conditions while the performance level of the adults is maintained. It would be of interest to investigate the performance of both younger and older children to determine whether this negative practice effect is a stable finding and whether it decreases with increasing age. The apparent similarity in performance of normal children and brain-damaged adults also raises the question as to whether brain-damaged and defective children would show this negative practice effect to an even more pronounced degree than that shown by normal children of comparable mental age.

The boys and girls showed virtually identical initial choice RTs (468 vs. 469 msec.). This finding contrasts with the indications of early studies that boys tend to show slightly faster RTs than girls. Since most of these studies were concerned with simple RT, it is possible that the sex difference is demonstrable in simple RT but not in choice RT. Another possibility is that the discrepancy in the findings of the present study, as compared with earlier investigations, is ascribable to sampling fluctuations. However, still another possibility is that the sex difference which was noted by earlier observers is actually disappearing as a function of changing conditions of nurture. More extensive observations on both simple and choice RT will be necessary to decide among these alternative explanations.

SUMMARY

The purpose of this study was to determine the effects of three types of motivating instructions on choice RT in school children. The sample consisted of 30 boys and 30 girls between the ages of 7 years, 7 months and

WAYNE A. OWEN

9 years, 4 months. Within the limits imposed by the maintenance of equality with respect to sex, the children were assigned randomly to the three instructional categories, viz., standard, urging, relaxing.

Initially, 30 visual choice RT trials were administered to each *S* under standard conditions. This was followed by differential instructions and by 30 postinstructional trials identical with the initial 30 trials. The criterion score for each *S* was the difference in mean time scores between his 30 preinstructional and 30 postinstructional trials.

The following results were obtained: (a) a significant main effect for instructions was found; (b) the standard instructional condition resulted in a significant loss in speed of performance as compared with preinstructional performance and differed significantly in its effect from the urging instructional condition but not from the relaxing instructional condition; (c) the urging instructional condition resulted in a significant gain in speed of performance as compared with preinstructional performance and was significantly different in its effect from each of the other two instructional conditions; (d) the relaxation instructional condition did not result in a significant change in speed of performance, as compared with preinstructional performance; (e) there was no significant difference between boys and girls in absolute level of initial choice RT performance.

REFERENCES

1. BENTON, A. L., & BLACKBURN, H. L. Practice effects in reaction time tasks in brain-injured patients. *J. abnorm. soc. Psychol.*, 1957, 54, 109-113.
2. BLACKBURN, H. L. Effects of motivating instructions on reaction time in cerebral disease. *J. abnorm. soc. Psychol.*, 1958, 56, 359-366.
3. BLACKBURN, H. L., & BENTON, A. L. Simple and choice reaction time in cerebral disease. *Conf. Neurol.*, 1955, 15, 327-338.
4. FARBER, I. E., & SPENCE, K. W. Effects of anxiety, stress, and task variables on reaction time. *J. Pers.*, 1956, 25, 1-18.
5. JOHANSON, A. M. The influence of incentive and punishment upon reaction time. *Arch. Psychol.*, 1922, No. 54.
6. ROSENBAUM, G. Stimulus generalization as a function of experimentally induced anxiety. *J. exp. Psychol.*, 1953, 45, 35-43.

an
Pa
om
lev
(2
th
be
th
sc
sig
an

stu
int

ch
an

the
an
Ro

CA

CHILDREN'S MANIFEST ANXIETY AND INTELLIGENCE¹

A. JACK HAFNER* and ARTHUR M. KAPLAN²

Washington University School of Medicine

A children's form of the manifest anxiety scale (CMAS) for fourth, fifth and sixth grade children has been reported by Castaneda, McCandless, and Palermo (1). They found that girls scored significantly higher than boys on the CMAS and L scale and found no significant relationship between level of anxiety and L scale scores. In a study by McCandless and Castaneda (2) a significant negative correlation was found between Otis IQ scores and the CMAS for sixth grade girls and a nonsignificant negative correlation between the same scores for sixth grade boys. Wirt and Broen (3), using the CMAS with fourth, fifth and sixth grade children, also found girls to score significantly higher than boys on the anxiety items, but found a non-significant relationship between Kuhlmann-Anderson or Binet IQ scores and the CMAS.

Because of the conflicting findings regarding intelligence, the present study sought to investigate further the relationship between the CMAS and intelligence.

METHOD

The CMAS was administered to 108 predominantly white³ fifth grade children of the St. Louis Public Schools who were from upper lower class and lower middle class families. This group of children was comprised of 53

* Department of Psychology, University of Minnesota, Minneapolis 14.

¹ The authors wish to express their appreciation for the assistance they received from the following persons: Dr. William Kottmeyer, Assistant Superintendent of Elementary and Special Education of St. Louis; Isabelle McGrath, Principal, Hamilton School; Ruth Rowe, Principal, Garfield School; and Shirley Millstone.

² Now at Eastern Maine Guidance Center, Bangor, Maine.

³ There were 3 Negroes in the group.

CHILD DEVELOPMENT

girls and 55 boys with a mean age of 132.1 months and a SD of 6.05 months who had been administered the Otis Beta three months previous to the CMAS administration. Their mean IQ was 103.8 with a SD of 10.62.

RESULTS

Table 1 presents the statistical results for the anxiety scale and L scale scores according to sex. No significant sex differences were found for the anxiety or L scale. The product-moment correlation between anxiety scale and L scale scores was .03.

TABLE 1
ANXIETY SCALE AND L SCALE MEANS AND SDs BY SEX

	ANXIETY SCALE				L SCALE			
	N	M	SD	t	N	M	SD	t
Boys	55	14.13	7.69	.51	55	4.11	7.25	.01
Girls	53	14.94	8.36		53	4.13	6.47	

The product-moment correlation between the Otis and CMAS was $-.21$ which was significant at the .05 level. The product-moment correlation between the Otis and L scale was $-.09$.

DISCUSSION

The lack of significant sex differences for the CMAS in the present study is contrary to previous findings and may be attributable to population differences. An explanation in terms of population differences must remain speculative, however, since the previous studies (1, 3) did not describe their groups any more specifically than in terms of sex and grade level. The possibility of population differences is suggested in the finding that the mean CMAS scores for both the boys and girls in the present study is lower than those found in the two previous studies (1, 3). It is also interesting to note, however, that the smallest sex difference for the CMAS in the Castaneda, McCandless, and Palermo study was with fifth grade children and the present investigation was confined to that grade. Wirt and Broen did not report a breakdown by grade level.

Population differences may also have been a factor in accounting for the present findings of a lack of sex differences for the L scale in contrast to the previously reported results (1). This is suggested in that the mean L scale scores for the boys and girls in the present study were higher than those of the previous study.

A. JACK HAFNER and ARTHUR M. KAPLAN

The results of the present study point toward a significant negative relationship between the CMAS and IQ as was also found by McCandless and Castaneda. A possible explanation for Wirt and Broen's lack of confirmation of this finding may have to do with the time between the CMAS and intelligence test administration. While there was only a three-month interval between testings in the present study and that of McCandless and Castaneda, some of the IQ scores for Wirt and Broen's study were obtained when the children were in kindergarten. This means that the CMAS scores of the Wirt and Broen study were correlated with IQ scores obtained from four to six years previously. The reliability of IQ scores over that long a period of time tends to decrease noticeably so the IQ scores may not have been very representative of the children's intellectual level at the time of the CMAS administration. Additional research should help clarify this issue of intelligence and the CMAS.

SUMMARY

The children's manifest anxiety scale was administered to a group of fifth grade children to determine its relationship with intelligence. A low but significant negative correlation was found to exist. No relationship was found between the L scale and intelligence and no sex differences were found in regard to anxiety or L scores.

REFERENCES

1. CASTANEDA, A., McCANDLESS, B. R., & PALERMO, D. S. The children's form of the manifest anxiety scale. *Child Developm.*, 1956, 27, 317-326.
2. McCANDLESS, B. R., & CASTANEDA, A. Anxiety in children, school achievement, and intelligence. *Child Developm.*, 1956, 27, 379-382.
3. WIRT, R. D., & BROEN, W. E. The relationship of the children's manifest anxiety scale to the concept of anxiety as used in the clinic. *J. consult. Psychol.*, 1956, 6, 482.

EFFECTS OF DELAYED REWARD PRETRAINING ON DISCRIMINATION LEARNING OF CHILDREN

LEWIS P. LIPSITT,* ALFRED CASTANEDA, and

JOAN DANSKIN KEMBLE¹

State University of Iowa

That immediate reward following the occurrence of an instrumental response is more effective than delayed reward in augmenting response strength has been reported by several investigators (2, 3, 5, 6) using infra-human Ss. In a recent study, Lipsitt and Castaneda (4) investigated choice behavior and speed of responding in young children as a function of delay of reward. In a within-Ss design, two stimulus lights were associated with differential delay of token-reward (marbles) following response to each. Response measures were percentage choice of the immediate-reward light on free-choice trials, and speeds to the alternate lights administered on forced-choice trials. The frequency of choice of the immediately rewarded stimulus over the delayed reward stimulus increased significantly over trials.

The present study is concerned with the effects of different periods of delay associated with each of two stimuli on *subsequent* performance in a simultaneous discrimination learning problem involving the same two stimuli. An experimental group is rewarded immediately when responding in the presence of one stimulus and after a delay when responding in the presence of the other stimulus. A control group experiences immediate reward to each of two stimuli. The two stimuli are subsequently employed as the relevant stimuli in a simultaneous discrimination task in which the stimulus previously associated with immediate reward for the experimental group is designated the correct stimulus for both groups. If the principal effect of pretraining is the acquisition of differential response strengths in favor of the stimulus associated with immediate reward, performance in the discrimination task should be facilitated for the experimental group.

METHOD

Subjects and Design

The Ss were 34 children from the Iowa Preschool Laboratories ranging in age from 3.1 to 5.5 years with a mean age of 4.8 years. They were assigned to one of two groups matched in pairs on the basis of age and

* Psychology Department, Brown University, Providence 12, Rhode Island.

¹ The writers are indebted to Mr. John Peterson for aid in constructing the apparatus.

CHILD DEVELOPMENT

sex. The experimental group experienced immediate reward to one color, and reward after a delay of 6 sec. to the other color, in the pretraining task. The color associated with immediate reward during pretraining was designated the correct stimulus in the discrimination task. The control group received immediate reward to both colors in pretraining, and the color designated correct in the discrimination task was the same as that for the experimental *S* of the matched pair. The red stimulus was correct for nine pairs and blue for eight pairs of *Ss*.

Apparatus

The *pretraining* apparatus consisted of a black plywood box 24 by 7 by 14 in. Two jewelled reflectors, one red and one blue, illuminated by 6-v. pilot lamps, were placed vertically and as close as possible to each other to the top center of the apparatus. The response element, a wooden circular disc fastened to a leaf-actuated microswitch, was situated approximately 2 in. below the lights. The lights and the response button were backed by a 6 in. square field of white paint. A plastic container, into which the token-rewards (marbles) were ejected from a copper tube inside the apparatus, was fastened 6 in. below the response button. On any given trial one of the stimulus lights and a Standard Electric Timer (situated to the back of the apparatus) were activated simultaneously by *E*. Depression of the button by *S* broke the circuit to the timer and simultaneously activated one of two electronic interval timers which controlled delivery of the marbles and offset of the light. One of the interval timers was set for immediate delivery of the marble (by solenoid action) while the other was set for a delay of 6 sec. A board containing two parallel rows of shallow holes of sufficient size to contain marbles was placed directly to the side of the apparatus in view and reach of *S*.

The *discrimination* apparatus was identical to the one just described with the exception that it contained two stimulus-response-reward units instead of one. These units were placed beside each other on the front panel 12 in. apart. Here the two colors could be presented simultaneously with position randomized, i.e., each color could appear on either side. Depression of the response button under the correct color immediately delivered a marble and turned off both colored lights. If the button under the incorrect color was depressed, both lights were turned off and a marble was not delivered. This arrangement permitted learning by the non-correction method. No delay of reward was involved in the discrimination task; the *S* was either rewarded immediately for a response or not at all, depending upon whether correct or incorrect.

Procedure

The procedure in *pretraining* was such that each color was presented 20 times in random order for a total of 40 trials. Upon completion of these 40 trials, the *S* was informed that the remainder of the marbles to fill the

board were to be obtained from the discrimination apparatus. The *discrimination* task was described as one in which the two colors would appear together with the colors appearing in different positions on different trials. This was demonstrated twice (with the colors in each of the two positions), and S was informed that if he pressed the button under the correct color he would obtain a marble. If the incorrect button was selected, a marble would not be delivered. The names of the colors were not spoken by E. The stimuli were presented until the remainder of the marble board was filled, requiring a minimum of 20 trials.

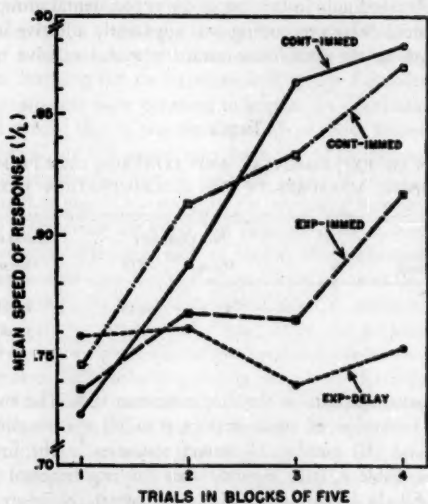


FIGURE 1—Response speeds to the two immediately rewarded stimuli for the control group, and for the immediate and delay stimuli for the experimental group, as a function of trials.

RESULTS

A reciprocal transformation of the pretraining data was performed on all latency scores and are presented, in blocks of five trials, for each light and the two groups separately in Figure 1. It can be noted that speeds to both lights are generally faster, after the first block of trials, for the control than the experimental group. There appear to be no consistent differences between speeds to the two lights (both immediately rewarded) for the control group. In the case of the experimental group, on the other hand,

CHILD DEVELOPMENT

after the first block of trials there is an increasing difference in speeds to the two lights in favor of the light associated with immediate reward. The tendency for speeds of the experimental group to the immediately rewarded light to be somewhat slower (nonsignificantly) in comparison to either of the two immediately rewarded lights for the control group may be due to the generalization of the effects of delay from the delayed reward light. Separate comparisons of the differences in speed to the two lights at the first block of trials for each group indicated that each group reacted with equal speeds to both lights. However, when the same comparisons were made at the last block of trials, a significant difference ($t = 2.86$, $df = 16$, $p < .02$) was obtained only in the case of the experimental group, indicating that the differential delay pretraining was apparently effective in increasing response strength to the immediate-reward stimulus relative to the delay stimulus.

TABLE I
COMPARISON OF EXPERIMENTAL AND CONTROL GROUPS FOR TWO
LEARNING MEASURES IN THE DISCRIMINATION TASK

Group	NO. CORRECT		TRIALS TO CRITERION	
	Mean	SD	Mean	SD
Experimental	14.9	4.45	27.71	8.00
Control	13.1	3.76	33.06	8.70

Table I presents the data of the discrimination task. The two measures used were: (a) number of trials necessary to fill the remainder of the marble board, and (b) number of correct responses in the first 20 trials. On the basis of Table I, it is apparent that the experimental group took fewer trials to fill the board and made more correct responses in the first 20 trials. For the number of trials to fill the board, the difference between the two groups (t test for related groups) was significant at the .001 level ($t = 4.59$, $df = 16$) while the difference in number of correct responses was significant at the .02 level ($t = 2.60$, $df = 16$).

DISCUSSION

The results based on the pretraining data of the present experiment support statistically a similar, though unreliable tendency in an earlier experiment (4) for faster speeds to be associated with immediate reward than delayed reward in children. One of the major procedural differences between the two experiments was the use of interspersed choice trials in the earlier study. Their omission in the present experiment may have facilitated the development of differential speeds.

The discrimination data are in accord with the hypothesis that the effect of differential delay pretraining is to produce differential response strengths which are transferred to the discrimination task, thus facilitating learning to respond to the previously immediately rewarded stimulus, now the correct discrimination stimulus. Other considerations, however, seem relevant to the interpretation of the present finding. One concerns the possibility that the principal effect of pretraining was to encourage the development of distinctive symbolic responses identifying or differentiating the delay periods associated with each stimulus. In a few instances, verbalizations of this type (e.g., "long," "short") did occur overtly. Such symbolic responses, according to Dollard and Miller (1), may serve to augment the distinctiveness of stimuli with a resultant decrease in generalization between them. In the present case, such a mechanism would serve to facilitate subsequent discrimination learning for the experimental group. Likewise, the obverse of this mechanism may have operated to impair the discrimination learning of the control group; that is, the equal periods of delay following responses to the two stimuli for this group may have served to increase the effective similarity of these already distinctive stimuli (red and blue).

Finally, another possibility concerns the orienting behavior of Ss in the present study. Increasing attention has recently been focused on the role played by receptor orienting acts in visual discrimination learning of inarticulate organisms (7, 8, 9). Factors affecting orienting behavior of young children are likely important determinants of performance in visual discrimination learning. It was noted that Ss in the present experimental group tended to develop more distinctive head and eye movements than did the control group. Such orienting habits may have developed as a result of the differential delay pretraining which the experimental group had received. It follows from this that, if such orienting habits facilitate discrimination learning, the learning of a group given differential delay pretraining with the previously *delayed* reward stimulus becoming correct in discrimination should *also* be facilitated. Such a prediction would be contrary to the expectation of impairment for such a group under the assumption that differential delay pretraining serves to develop differential response strengths to the stimuli which transfer directly to the discrimination task.

SUMMARY

Two groups of 17 preschool children, matched for age and sex, received 20 rewarded trials to each of two distinct colors. The experimental Ss received immediate reward to one color, and reward after a delay to the other. Reward was immediate to both colors for the control Ss. There was no significant difference in speed of response to the two immediately rewarded stimuli in the control group. A significant difference in response speed to the delayed and immediately rewarded stimuli, however, was found in the experimental group at the end of pretraining.

CHILD DEVELOPMENT

All Ss were subsequently tested in a color-discrimination task, in which the stimulus associated with immediate reward (for the experimental Ss) in the pretraining task was now correct for all Ss. Ss learned by the non-correction method and were immediately rewarded for correct responses. Experimental Ss took significantly fewer trials to reach criterion and made significantly more correct responses in the first 20 trials.

Differential delayed reward pretraining apparently serves to facilitate learning in a transfer task involving the discrimination of the pretraining stimuli. Several mechanisms possibly accounting for this effect were discussed.

REFERENCES

1. DOLLARD, J., & MILLER, N. E. *Personality and psychotherapy*. New York: McGraw-Hill, 1950.
2. GRICE, G. R. The relation of secondary reinforcement to delayed reward in visual discrimination learning. *J. exp. Psychol.*, 1948, 38, 1-16.
3. HARKER, G. S. Delay of reward and performance of an instrumental response. *J. exp. Psychol.*, 1956, 51, 303-310.
4. LIPSITT, L. P., & CASTANEDA, A. Effects of delayed reward on choice behavior and response speeds in children. *J. comp. physiol. Psychol.*, 1958, 51, 65-67.
5. LOGAN, F. A. The role of delay in reinforcement in determining reaction potential. *J. exp. Psychol.*, 1952, 46, 393-399.
6. PERIN, C. T. The effect of delayed reinforcement upon the differentiation of bar responses in white rats. *J. exp. Psychol.*, 1943, 32, 95-109.
7. PUBOLS, B. H., JR. The facilitation of visual and spatial discrimination reversal by overlearning. *J. comp. physiol. Psychol.*, 1956, 49, 243-248.
8. REID, L. S. The development of noncontinuity behavior through continuity learning. *J. exp. Psychol.*, 1953, 46, 107-112.
9. WYCKOFF, L. B. The role of observing responses in discrimination learning. *Psychol. Rev.*, 1952, 59, 431-442.

AN EVALUATION OF THE CHANNEL SYSTEM ON THE BABY GRID

DAVID SMILLIE*¹

The Merrill-Palmer School

Of the wide variety of growth assessment standards developed in the past 25 years, the Wetzel Grid (5) has received widespread attention in the literature. However, Wetzel's Baby Grid (7), which extends the Grid technique to the period of infancy, has received little descriptive or critical attention (1). The present paper describes and evaluates the Baby Grid with reference to 230 infants seen at The Merrill-Palmer School in Detroit.

The Baby Grid is essentially a chart providing two complementary methods for assessing changes in length and weight during infancy. These methods are used to provide a measure of healthy progress through the evaluation of physique constancy and change in size. The "channel" portion of the five-color Baby Grid consists of a graph on which length is plotted on the abscissa and weight on the ordinate. Increases in length and weight are represented by a line running diagonally from lower left to upper right. A series of lines running parallel to the "usual" progress of an infant are called "channels." The channel lines are crossed at regular intervals by another system of lines called "developmental levels." The channel lines represent guide lines for the maintenance of a consistent physique; the crossing of levels indicates an increase in size.

* The Merrill-Palmer School, 71 Ferry Avenue, East, Detroit 2, Michigan.

¹ I wish to thank the many persons at Merrill-Palmer who have helped to collect and interpret the materials used in this study. In particular I wish to express my gratitude to Dorothy Tyler, Editor of Publications, for her critical reading and editing of this paper.

It should be noted that Wetzel in his paper on the Baby Grid thanks Mary E. Sweeney and Marian E. Breckenridge of The Merrill-Palmer School "for placing many of their long-term careful records at his disposal. The final stages of design and construction . . . were greatly expedited by this opportunity to study their data objectively and then to compare the results with their personal observations and histories of those infants and children" (7, p. 454).

CHILD DEVELOPMENT

The second method for assessing length and weight changes consists of a series of lines ("auxodromes") placed to the right of the channel system. These lines represent percentile curves of the size factor, as measured by the developmental levels, against age. This portion of the Baby Grid provides a way of comparing an individual infant's size with that of other infants of the same age.

The Baby Grid also provides panels for graphing chest and head circumference and for determining caloric allowances for any given size. Immediately to the right of the channel system, suggested levels for achieving motor abilities are given.

The findings in the present paper concern the channel system.

PRINCIPLE OF THE GRID TECHNIQUE

The principle on which Wetzel bases his Grid technique is that an individual growing in a healthy manner will maintain, during most of his growing years, a constant ratio of height to weight, as plotted on a double logarithmic scale. This constancy of height-weight ratio is represented on both Grids as straight-line progression along a given channel. That is, an individual who is maintaining healthy growth progress is expected to stay within one channel. Which channel is followed is assumed to depend on his inherent physique pattern.

According to Wetzel, there is one period in growth progress from birth to adulthood when the individual does not follow this straight-line progression. This period occurs between 1 year and 6 years, when the chubby physique of the infant changes to the more linear physique of the child. This shift, which brings the Baby Grid system in line with the channels of the Grid for children, is represented on the Baby Grid by a yellow band, slightly over two channels in width. This yellow band, curving to the right, away from the straight-line channels, then becomes the frame of reference for healthy growth.

Wetzel (7) says the results of plotting length and weight on the Baby Grid "are practically self-evident, since the significance of the original measurements is directly revealed in terms of the fidelity with which the curves of these different infants conform to certain obvious patterns of the Grid"² (p. 442). Again: "Ordinarily, shifts of one channel in 20 levels of forward advancement are acceptable, but even these should not persist. . . . On the turns, a baby's curve should follow the trend of the shaded pink and yellow path with corresponding certainty . . . deviations outside these limits indicate, apart from frank errors, that the departures should be considered the result of trouble, and hence that the babies who fail this test of quality should be examined for cause" (p. 449). These points are

² In giving a random sampling of 25 Baby Grids to three staff members familiar with its use, it was evident that they interpreted the results with regard to abnormality differently even when given Wetzel's criteria.

DAVID SMILLIE

illustrated with cases of feeding difficulties, overfeeding, and clinical pathologies.

A number of questions arise from Wetzel's presentation of the Baby Grid. Some can be answered by an account of how the plotting of data from a sample of healthy babies actually appears on the Baby Grid; these questions are discussed in terms of the data presented in this paper. However, several difficulties emerge from the use of the Baby Grid in relation to Wetzel's claims for it. One such difficulty arises from the ambiguities of interpretation. For example, when Wetzel gives, as a criterion of lack of health, deviations greater than one channel in 20 levels, he does not point out that it may take an infant a full 12 months to traverse 20 levels at the upper limits of the Baby Grid, while in the neonatal period an infant may traverse 20 levels in a period of 2 weeks. Is one to infer that these two extremes should be interpreted in the same manner? Is it necessary to wait the full 12 months before pronouncing difficulty at the upper age limits of the Baby Grid?

Furthermore, the interpretation of a shift of one channel for 20 levels, somewhat difficult during the straight-line portion of the Baby Grid, is next to impossible to make when the individual shifts away from the channels entirely at about 10 months of age. During this period dotted lines parallel to the yellow band presumably indicate expected curvature for the extremes of physique, but during this period of curvature there is no means of referring to channels.

NATURE OF QUESTIONS DISCUSSED

Thus, questions about the Baby Grid discussed in this paper are of two kinds. The first kind are those that may be answered as a result of testing it continuously with accumulated longitudinal data. For example, using the criteria of Wetzel, what percentage of infants satisfy the requirement for healthy channel progress? Are there sex differences? If so, do they parallel those of the Grid for children, with more females than males deviating (2)? Is the percentage of deviations similar from age to age? Does the distribution of physiques as measured by the Baby Grid follow a Gaussian curve, or is it, like the Grid, skewed in the direction of the stockier physiques (3)?

In addition, since the Baby Grid introduces a principle, a period of curvature, not embodied in the Grid, another kind of question arises, namely: To what extent does curvature occur within well defined limits? Can patterns of "normal" curvature be described? To what extent is curvature influenced by such factors as size-for-age, sex, and physique? These questions have several considerations as their basis. Thus, when one considers a straight line only, any deviation from that line becomes a sign of departure from healthy progress; nor does a straight line offer a measure of early or late development or of what Wetzel calls "speed." (The auxodromal curves to the right of the channel system are for this purpose.)

CHILD DEVELOPMENT

However, on the Baby Grid there is a measure of "speed" not necessarily related to size, and this is *change* of physique. There is, however, no way of differentiating a normal shift to the right (beginning curvature) from a pathological shift (change of physique due to illness), unless the normal shift falls within narrowly defined limits.

SUBJECTS

The population of 230 infants used in this study is drawn from the records collected at The Merrill-Palmer School over the past 25 years. The babies were seen in the Infant Service as a part of the well-baby program. Weight measurements were taken on a standard infant scale. Recumbent length measurements were made with a measuring board. Infants were weighed and measured at varying intervals, from once a month to once in 6 months over a period of 2 years. In most cases two anthropometrists measured the infants, coming to an agreement of one eighth of an inch of each other. In a few cases one person made the measurements.

The parents of the babies studied were primarily of Old American and North European stock. The parents were 95 per cent native born; for 50 per cent of the infants, all four grandparents were native born. The parents in general had professional and business occupations. Both mothers and fathers had a high level of education. Over 98 per cent of the parents had

TABLE I
WEIGHTS AND LENGTHS OF INFANTS IN MERRILL-PALMER SERIES

Age (months)	MALES			FEMALES		
	N	M	SD	N	M	SD
WEIGHT (POUNDS)						
3	74	13.85	1.44	69	12.45	1.39
6	105	17.84	1.84	111	16.62	1.61
9	93	21.00	2.12	86	19.28	2.09
12	100	22.71	2.37	95	21.54	2.22
18	82	25.84	2.32	86	24.34	2.37
24	80	27.83	2.60	78	26.96	2.57
LENGTH (INCHES)						
3	73	23.93	.88	67	23.46	.60
6	106	26.51	1.00	113	26.01	.78
9	91	28.37	1.06	82	27.71	1.01
12	102	29.75	.91	94	29.28	.99
18	80	32.30	1.13	81	31.91	1.26
24	76	34.43	1.10	75	34.19	1.20

DAVID SMILLIE

some college education, and in 40 per cent of the families one parent had a graduate degree.³

The infants of the present study are therefore selected from a rather homogeneous population. Since some of the findings do not show a close relationship to those represented on the Baby Grid, a table of cross-sectional values is presented, reporting means and standard deviations for weight and length for males and females from the Merrill-Palmer series (Table 1).

PROCEDURES AND CRITERIA

In classifying the Baby Grids of the Merrill-Palmer series, the following procedures and criteria were used. First, the angulation of the lines connecting successive plots for each child was measured with a protractor to obtain an accurate quantitative measure of inclination. Perfect channel conformance during the straight-line portion of the Baby Grid was found to be represented by a line having a slope of 125° , while a line crossing one channel in 20 levels to the left had a slope of 116° and to the right, 134° . To account for Wetzel's statement that shifts of one channel in 20 levels should not persist, a changing scale was used for lines exceeding 20 levels in length, with criteria as follows:

CRITERIA FOR NONCONFORMANCE TO CHANNELS

Straight-line Portion

Distance (levels)	Angulation Limits (degrees)
-20	No readings made
20-30	Below 116; above 134
30-40	Below 117; above 133
40-50	Below 118; above 132
50+	Below 119; above 131

Curved Portion

Curve should not deviate more than the width of a channel in its relation to the curving yellow band. Greater deviations, compensated for after a short span, are not included as nonconforming.

Examination of the deviations from the curved portion showed that many curves that did not fit the normal pattern failed to satisfy the criterion, not because they were of the wrong shape, but because the shift to the right took place at a level inappropriate according to the chart. By tracing the curve on a transparency and moving it up or down along the channels, it was found that a number of infants fell exactly on the curve represented by the yellow band. These infants were classified as "early curvers" if the trans-

³ These figures are based on approximately half the total sample, since the relevant information is lacking in the other cases.

CHILD DEVELOPMENT

parency had to be moved up to achieve a fit, and "late curvers" if it had to be moved down.

FINDINGS

Table 2 shows the number and percentage of infants deviating from or conforming to "normality" on the Baby Grid. According to our interpretation of Wetzel's standards, a large proportion do deviate at one time or another during the first 2 years. A number of factors enter into this high proportion of deviations. Thus, 15 per cent are classified as deviating because they started the thinning-out process (curvature) at a level below or above the expected point—a problem later discussed in further detail. Others probably deviate because of undetected inaccuracies in measuring or plotting. Still others may deviate because of temporary illnesses, digestive upsets, and feeding problems, which are to be expected in any group of infants. While all these factors may contribute to the high percentage of deviant individuals, they are not special conditions of this study but would apply to the use of the Baby Grid in other situations.

TABLE 2
MERRILL-PALMER INFANTS DEVIATING FROM AND
CONFORMING TO "NORMALITY" ON
THE BABY GRID ($N = 230$)

<i>Infants</i>	<i>N</i>	<i>Per Cent</i>
Deviating	169	73
Conforming	61	27

In examining more closely the total number of deviations (rather than the number of individuals who deviate), they are seen to break down into 40 per cent during the straight-line portion of the Baby Grid and 60 per cent during the curved portion. This percentage corresponds almost exactly to the proportional periods of time covered by the straight-line and curved portions. From the point of view of time, therefore, we may say that there are equal numbers of deviations in the curved and straight-line portions.

In Table 3 deviations have been classified according to type and degree (major and minor). This classification involves some subjectivity, for it is based on inspection rather than on clinical investigation of case histories or clinical experience with the Baby Grid. In general, major deviations are characterized by sudden and sustained shifts of more than two channels and by persistent falling off toward the slender channels. However, some excep-

TABLE 3

INCIDENCE OF DEVIATIONS BY TYPE AND DEGREE

Item	Deviations	N	Per Cent of Total Deviations
MINOR DEVIATIONS			
1. A deviation of the birth position from the prevailing trend	41	17	
2. One point out of line with the prevailing trend	20	8	
3. A shift in the prevailing trend at a single point	18	7	
4. A shift toward a stocky physique with a return to the original channel	14	6	
5. A minor deviation during the curved portion of the Baby Grid	42	17	
6. Curvature takes place earlier than expected (curve early)	55	23	
7. Curvature takes place later than expected (curve late)	13	5	
MAJOR DEVIATIONS			
8. Major deviations during the straight-line portion of the Baby Grid ..	16	7	
9. Major deviations during the curved portion of the Baby Grid	23	9	

tions should be noted. Thus, items 1 and 6, classified as minor deviations, for reasons explained later in this paper, may involve either sudden shifts of more than two channels or, as in item 6, a persistent falling off toward the slender channels. According to this scheme, only 16 per cent of deviations are major in degree. However, the classification involves a judgment made in each case on the basis of all available measurements, and, until successive points are plotted, any deviation from the standard is potentially of major degree. Thus, for one who assesses growth progress on the Baby Grid from month to month, a point eventually seen to represent the beginning of curvature before it is expected, as in item 6 of Table 3, may seem to be a major deviation toward the slender channels.

Table 3 shows most of the deviations to be minor, with a large percentage representing a discrepancy between the channel position at birth and the subsequent prevailing trend. The measurements indicate that between birth and 1, 2, and 3 months of age, 41 per cent of cases deviate, according to our interpretation of Wetzel's standards, while from 3 months to 4, 5, and 6 months only 23 per cent of cases deviate. Whether this difference is attributable to inaccuracy of birth measurements (taken mostly at the hospital), to individual differences in postnatal adjustment,⁴ or to the logarithmic nature of the graph, which is more sensitive to variations in height and weight at the lower levels, is not clear. However, the findings suggest that deviations from the Baby Grid during the first several months of life should be interpreted leniently.

⁴ In this regard, see Tanner's (4) interesting hypothesis of differential response to prenatal environment.

CHILD DEVELOPMENT

The meaning of the minor deviations during the curved portion of the Baby Grid is also rather difficult to interpret. Wetzel (6) implies that the Grid for children is useful only when the child settles in a given channel, and that the thinning-out period is unpredictable. Whether the *beginning* of the thinning-out period shown on the Baby Grid is also unpredictable, he does not state.

The highest number of deviations occurs in cases in which the thinning-out process (i.e., curvature) starts earlier than expected (Table 3, item 6). In some cases (item 7), the curvature occurs later than expected. This finding concerning early and late "curvers" suggests that, in terms of Wetzel's formula, there is a "normal" phenomenon not adequately described by the Baby Grid.

Many infants described here as early and late curvers follow the Baby Grid curvature point for point, but at a level which must be considered "abnormal" according to Wetzel's standards. Thus, the late curvers follow the straight-line channels beyond the point where a shift to the right would be expected, while the early curvers shift to the right earlier than would be expected, cutting across the channels. Thinning out at, say, level 60 may indicate a deviation from normal progress and hence be a sign of incipient pathology; or it may represent a normal curve for the individual infant. The pattern on the channel portion of the Baby Grid offers no way to distinguish between these two phenomena.

There is no significant sex difference in the incidence of deviations, though there is a somewhat higher percentage of males (61.5) in the "curve late" group and of females (52.4) in the "curve early" group.

A more thorough analysis of thinning out in infancy is needed, since for the average infant it concerns slightly more than half the period from birth to 2 years. According to Wetzel (7), the "turn . . . is reached in about 11 to 12 months by average babies" (p. 442). According to the Baby Grid itself, the turn seems to begin at about level 50, at an average age of 9½ months. In the Merrill-Palmer infants the beginning of curvature occurs on the average at level 43, at age 10 months.

Factors Affecting Curvature

The only allowance made for variability in point of curvature on the Baby Grid is on the basis of physique; that is, infants in the heavier channels curve at a point somewhat earlier than infants in the slender channels. In order to test the relationship between the developmental level at which curvature begins on the Baby Grid with the three variables of physique, sex, and size-for-age,⁵ an analysis of variance was made. Significant relationships, at the 1 per cent level, were found to exist between level of curvature and physique, with stockier infants curving at an earlier developmental

⁵ "Physique" refers to the channel position of the infant at the point at which curvature begins; "size-for-age" refers to the age in months at which the infant reaches developmental level 50.

DAVID SMILLIE

level than slender infants; and between level of curvature and size-for-age, with large infants curving at a later developmental level than small ones. No significant relationship was found to exist between level of curvature and sex.

These findings indicate that the relationship between physique and point of curvature suggested on the Baby Grid does obtain, so that infants who are in the heavy channels at the time of thinning out start their curvature somewhat earlier than infants in the slender channels. In addition, the size of the individual infant (size-for-age), represented on the auxodromal side of the Baby Grid, provides a certain amount of predictive information in assessing the point at which curvature will begin, with large infants (advanced auxodromal position) curving at advanced developmental levels.

TABLE 4
COMPARISON OF LATE CURVERS, EARLY CURVERS, AND TOTAL

	<i>N</i>	<i>Males (%)</i>	<i>Size-for-Age*</i>	<i>Physique</i>	<i>Age of Curvature</i>
Late curvers	13	61.5	7.6	10.1	11.0
Early curvers	61	42.6	11.0	9.4	10.0
Total	242	51	9.2	9.5	10.3

* Chronological age at developmental level 50.

Table 4 shows the data from another point of view, namely, the characteristics of children who curve early and those who curve late. Here again, size-for-age is the salient characteristic; that is, late curvers are large for their age, early curvers are small, while sex and physique are not distinctively related to the phenomenon of curvature.

SUMMARY

Data are presented on the performance of 230 infants in the Merrill-Palmer series on the Baby Grid (Wetzel), applicable from birth to age 2 years. Findings were as follows:

1. Some deviation from the standards described by Wetzel was found in 73 per cent of the infants over the 2-year span.
2. Of these deviations, 83 per cent were of a minor nature, while 16 per cent were major.
3. The highest percentage of deviations were due to curvature (thinning out) earlier than expected.
4. A relatively large proportion of deviations occurred during the period from birth to 3 months.

CHILD DEVELOPMENT

5. Sex differences were not significant in either incidence or type of deviations.

6. Sex was found to be unrelated to the level of beginning curvature on the Baby Grid; physique to be related in the expected direction; while size-for-age was found to be directly related, with large infants curving at a later developmental level than small ones.

7. Size-for-age is the only characteristic of early and late curvers definitely related to the timing of curvature.

REFERENCES

1. BRECKENRIDGE, M. E., & MURPHY, M. N. *Rand, Sweeny & Vincent's Growth and development of the young child.* (6th Ed.) Philadelphia: Saunders, 1958.
2. SNODGRASSE, R. M. Growth in Chicago private school boys with special reference to the Wetzel grid. Unpublished doctoral dissertation. Univer. of Chicago, 1950.
3. TANNER, J. M. The assessment of growth and development in children. *Arch. Dis. Childh.*, 1952, 27, 10-33.
4. TANNER, J. M., *et al.* Aberdeen growth study: I. The prediction of adult measurements from measurements taken each year from birth to 5 years. *Arch. Dis. Childh.*, 1956, 31, 372-381.
5. WETZEL, N. C. Physical fitness in terms of physique development and basal metabolism. *J. Amer. med. Ass.*, 1941, 116, 1187-1195.
6. WETZEL, N. C. Growth. In O. Glasser (Ed.), *Medical physics*. Chicago: Year Book Publishers, 1944. Pp. 513-569.
7. WETZEL, N. C. The baby grid. *J. Pediat.*, 1946, 29, 439-454.

CHILDREN'S ACCEPTANCE OF RECIPROCITY AS A JUSTICE-PRINCIPLE

DOLORES DURKIN*

University of California, Berkeley

As reported in *The Moral Judgment of the Child*, Piaget (3) presented a variety of problematical situations to 167 subjects "from the poorer parts of Geneva (Switzerland)" and ranging in age from 6 to 12. On the basis of their responses he proposed a theory concerning the development of children's moral judgments. A significant element in his theory is the increasing acceptance, with age, of reciprocity as a justice-principle.

One section in the Piaget study, "Justice between Children" (3, pp. 294-312), is especially designed to examine this "eye for an eye" concept of justice. Subjects' responses to two problems are described. The first of the problems depicts a situation wherein a small boy, unable to return the blows received from a larger boy, hides the aggressor's apple and roll in a cupboard. Subjects are asked, "What do you think of that?" The second problem simply asks about what should be done "if someone punches you."

To both of these questions, Piaget points out in his summary analysis, the older subjects tended to respond in ways that reflected acceptance of reciprocity. For example, in the first situation the little boy, being unable to give back what he had received, "should be allowed to restore the balance of things by hiding the aggressor's lunch" (3, p. 301). In the second situation, as far as these older subjects were concerned, "it is strictly fair to give back the blows one has received" (3, p. 301).

Piaget's younger subjects, to the contrary, tended to condemn reciprocity as being forbidden, and they relied on the adult to restore right order in both situations described to them.

The two quite different attitudes adopted by Piaget's subjects, coupled with the problematical situations eliciting them, suggest at least three problems for consideration:

* School of Education, University of California, Berkeley 4.

CHILD DEVELOPMENT

1. If subjects who were similar in chronological age but different in nationality and socioeconomic class from those interrogated by Piaget were also questioned about the restoration of right order in instances of physical aggression, would their responses tend to show, as Piaget's theory proposes, that "children maintain with a conviction that grows with their years that it is strictly fair to give back the blows one has received" (3, p. 301)?

2. If the intelligence, as well as the chronological age of these subjects were considered, might it not be found that particular kinds of justice-concepts are associated with levels of intelligence as well as with chronological age?

3. If these same subjects were questioned about violations of justice other than the one of physical aggression, would their responses tend to reflect a developmental trend similar to the one described by Piaget?

The first of the problems outlined above was considered in an earlier study (1). The present investigator questioned 101 middle-class American children in grades 2, 5, and 8 about acts of physical aggression, and found that between grades 2 and 5 acceptance of reciprocity as a justice-principle did increase, but between grades 5 and 8 it decreased significantly. The older of the subjects, as well as the younger, tended to seek justice in the authority person.

The second problem outlined above was also considered in the writer's earlier study (1). Data for the second and eighth grade groups showed no significant relationship between a child's level of intelligence and his particular understanding of what is just. For the fifth grade group, however, the relationship was significant. Because of this lack of consistency in the findings, the problem of the function of intelligence in the development of justice-concepts will be considered again in the present study.

The major purpose of the present study, however, was to consider the third problem outlined above, that suggested by the limited scope of the justice-violations depicted in Piaget's stories. By enlarging, in kind, the violations about which subjects are questioned, the writer hoped to evaluate still further the general applicability of Piaget's theory concerning the increasing acceptance, with age, of reciprocity as a justice-principle.

SUBJECTS AND PROCEDURE

The 101 subjects in this study were all of the second, fifth, and eighth graders in a Midwestern community-consolidated school. Twenty-eight of the total group were in Grade II, 38 in Grade V, and the remaining 35 were in Grade VIII. The mean chronological age, in years, for the three groups was 7.8, 10.9, and 13.9.

For the purposes of this investigation, justice was operationally defined in terms of (a) the rendering of rights regarding one's property and (b) the rendering of rights regarding one's character. Possible violations of these two kinds of rights were depicted in story situations and subjects,

DOLORES DURKIN

in individual, tape-recorded interviews, were questioned about them. Identical story situations were presented to both boys and girls, but for the boys story characters were always male; for the girls, they were always female.

Following are the stories and questions (as presented to male subjects) designed to examine children's understanding of what is just in the area of property rights:

I

One morning in school Anton asked Cyril Hayes, the boy who sits in front of him, if he could use his eraser. Cyril said, "No." About a week later Cyril was the one who didn't have an eraser, so he asked Anton if he could use his.

What should Anton do?

Why?

II

In another room a boy by the name of Keith took a ruler off Russell Holec's desk and wouldn't give it back to him.

What should Russell do?

Why?

III

What if Keith had taken the ruler and broken it in two?

What should Russell have done then?

Why?

Following are stories concerned with defamation of one's character. The first is included for the purpose of establishing possible grounds for "vengeance"; responses to it are not analyzed in this study. The second story is designed to examine children's understanding of what is just in the area of character rights.

IV

One morning in school Basil copied answers from another boy's test paper, and his teacher caught him. Later when the children were out on the playground at recess time, Alf, a boy in Basil's room, ran up to Basil and whispered, "Hi, Cheater! Hi Cheater!"

What should Basil do?

Why?

V

On another day Alf was the one who was caught cheating.

What should Basil do?

Why?

After all of the interviews had been completed, the investigator administered the Kuhlmann-Anderson Intelligence Test to 100 of the subjects.¹ The range of IQ scores in grades 2, 5, and 8 was, respectively, 92-114,

¹ One second grader transferred before the test was administered.

CHILD DEVELOPMENT

69-122, and 77-148. The mean scores were found to be 103.4, 101.6, and 103.4. Standard deviations were 5.1, 13.2, and 11.4

FINDINGS

Responses and Grade Level

For purposes of this investigation, subjects' responses to all four stories were divided into two general categories: those that reflected acceptance of reciprocity and those that did not. The frequency distribution of these two kinds of responses at the three different grade levels is summarized in Table 1.

TABLE I
FREQUENCY DISTRIBUTION OF KINDS OF RESPONSES AT THE THREE
DIFFERENT GRADE LEVELS

Kind of Response	G R A D E			<i>h</i> *	Significance Level of Chi Square Value
	2	5	8		
<i>Sharing Property</i>				-.66	< .01
Reciprocity	20	15	6		
Nonreciprocity	8	23	29		
<i>Taking Property</i>				-.29	> .05
Reciprocity	2	2	1		
Nonreciprocity	26	36	34		
<i>Destroying Property</i>				-.55	> .05
Reciprocity	3	4	0		
Nonreciprocity	25	34	35		
<i>Defaming Character</i>				-.70	< .01
Reciprocity	14	8	2		
Nonreciprocity	14	30	33		

* Index of order association.

To measure the degree of association between age, as defined by grade level, and acceptance of reciprocity, the index of order association, denoted by *h*, was used (2, 4). The values of *h* for the four stories are included in Table 1. Since, to the writer's knowledge, a standard error formula has not yet been developed for *h*, the significance levels of chi square values were used to test hypotheses of zero correlation. These also appear in Table 1.

As the table shows, the tendency to accept reciprocity, contrary to Piaget's proposal, appears to be negatively related to chronological age. The relationship is statistically significant for two of the four stories used in this study.

A second and more specific categorization of responses showed, for story I (sharing property), that there were still just two kinds of solutions proposed by the subjects. As was pointed out above, one reflected acceptance

DOLORES DURKIN

TABLE 2

FREQUENCY DISTRIBUTION OF THE TWO KINDS OF RESPONSES ON THE BASIS OF THE IQ LEVEL OF THE RESPONDENT

IQ Level	KIND OF RESPONSE		Chi Square Value	Significance Level
	Reciprocity	Nonreciprocity		
<i>Grade II</i>				
Sharing Property			0.54	> .30
Above Median	10	3		
Below Median	9	5		
Taking Property			0.04	> .80
Above Median	1	12		
Below Median	1	13		
Destroying Property			0.50	> .30
Above Median	2	11		
Below Median	1	13		
Defaming Character			0.98	> .30
Above Median	8	5		
Below Median	6	8		
<i>Grade V</i>				
Sharing Property			0.10	> .70
Above Median	7	12		
Below Median	8	11		
Taking Property			2.12	> .10
Above Median	0	19		
Below Median	2	17		
Destroying Property			2.08	> .10
Above Median	1	18		
Below Median	4	17		
Defaming Character			0.64	> .30
Above Median	3	16		
Below Median	5	14		
<i>Grade VIII</i>				
Sharing Property			3.52	> .05
Above Median	1	17		
Below Median	5	12		
Taking Property			1.12	> .20
Above Median	0	18		
Below Median	1	16		
Destroying Property			0.03	> .80
Above Median	0	18		
Below Median	0	17		
Defaming Character			2.28	> .10
Above Median	0	18		
Below Median	2	15		
<i>Grades II, V, VIII</i>				
Sharing Property			1.50	> .20
Above Median	33	17		
Below Median	27	23		
Taking Property			0.21	> .50
Above Median	2	48		
Below Median	3	47		
Destroying Property			0.15	> .50
Above Median	3	47		
Below Median	4	46		
Defaming Character			0.00	> .99
Above Median	12	38		
Below Median	12	38		

CHILD DEVELOPMENT

of reciprocity; that is, it recommended that property should not be shared with a child who has previously refused to share his. The second kind of response recommended sharing with another regardless of his past behavior.

To stories II and III (taking property and destroying property), subjects' responses were more varied. The most common kind of solution proposed at all three grade levels was one which recommended telling an authority person (the teacher) about the child who has taken and destroyed another's property. This was given by 75.0 per cent of the second graders, 81.6 per cent of the fifth graders, and 85.7 per cent of the eighth graders. The other kinds of responses to stories II and III recommended verbal chastisement, forgiveness, and reciprocity.

Subjects' responses to story V (defaming character) were also of just two kinds. One kind condemned reciprocity by taking the position that no one ever has the right to defame another's character and, therefore, the offended boy in the story should not do or say anything. The second kind of response proposed reciprocity in that it maintained that, whenever one's character has been defamed by another, it is just and right, given the opportunity, to defame his character in return.

Responses and IQ Level

Table 2 shows the frequency distribution of reciprocity and nonreciprocity responses on the basis of the IQ level of the respondent. It also lists the values of chi square when hypotheses of no relationship between kind of response and IQ level were tested. Again using the 5 per cent level as the criterion of significance, it appears that kind of response, in terms of its being "reciprocity" or "nonreciprocity," is not significantly related to the respondent's level of intelligence.

DISCUSSION

At least two kinds of comparisons, as has been pointed out, become relevant in an evaluation of the validity of that part of the Piaget theory which proposes an increasing acceptance, with age, of reciprocity as a justice-principle. The first is one which would compare the judgments of two culturally different groups of subjects about similar violations of justice; the second is one which would compare judgments of these same two groups of subjects about different kinds of justice-violations.

The first of these comparisons was made in the writer's earlier study (1) in which it was found that the trend in reciprocity responses for an American group of subjects was significantly different from that of Piaget's Swiss subjects even though the two groups were comparable in chronological age range, and were questioned about a similar kind of justice violation, namely, physical aggression. As was pointed out in that study, such a difference in findings would at least suggest that the Piaget theory minimizes the influence of the environment on a child's understanding of what is just.

DOLORES DURKIN

Perhaps, then, it is not valid to assert, as does Piaget, that "the sense of justice, though naturally capable of being reinforced by the precepts and the practical example of the adult, is largely independent of these influences . . ." (3, p. 195).

The second kind of comparison suggested above has been made in the present study. Its findings show, first of all, that when justice-violations are depicted in terms of violations of property rights and of character rights, the tendency for subjects to accept reciprocity as a principle of justice is negatively related to their chronological age. In this study, contrary to what Piaget's theory would suggest, it was the second graders rather than the eighth graders who most frequently tended to see restoration of right order in terms of reciprocity.

While a comparison of responses within this single group of subjects was not the concern of this study, its findings nonetheless show (*see* Table 1) that an identical group of subjects can make quite different kinds of judgments about different kinds of justice-violations. This variation in response becomes even more evident when the same kind of comparison is made on an individual rather than a group basis. In this study, for example, only four subjects out of the total group of 101 consistently proposed reciprocity for all four stories. This phenomenon of specificity suggests a further flaw in the Piaget theory in that the theory purports to describe the development of children's attitudes toward reciprocity on the basis of judgments about a single kind of justice-violation, that of physical aggression.

In summary, all of these various kinds of comparisons would tend to indicate that the factors affecting a child's understanding of justice are sufficiently multiple and that justice, operationally defined, is sufficiently complex that any theory which attempts to explain "The Development of the Idea of Justice in Children" is, from the start, doomed to inevitable overgeneralization and consequent error.

SUMMARY AND CONCLUSIONS

The major purpose of this study was to evaluate the validity of Piaget's proposal concerning children's increasing acceptance, with age, of reciprocity as a justice-principle. Piaget's subjects, of various ages, were questioned about story situations which depicted violations of justice in terms of physical aggression; subjects in this study, also varying in age, were questioned about story situations which depicted justice-violations in terms of violations of property rights and of character rights. The distribution of reciprocity responses, on the basis of age, for the two different groups of subjects was then compared.

A secondary purpose of this study was to examine the relationship between a child's level of intelligence and his acceptance of reciprocity.

Findings in the study showed that:

CHILD DEVELOPMENT

1. Acceptance of reciprocity as a justice-principle, contrary to Piaget's proposal, decreased as the chronological age of subjects increased.

2. Acceptance of reciprocity did not appear to be related to a child's level of intelligence. This lack of relationship was consistent throughout the different age levels.

REFERENCES

1. DURKIN, DOLORES. Children's concepts of justice: a comparison with the Piaget data. *Child Developm.*, 1959, 30, 59-67.
2. GOODMAN, L. A., & KRUSKAL, W. H. Measures of association for cross classifications. *J. Amer. stat. Ass.*, 1954, 49, 732-764.
3. PIAGET, J. *The moral judgment of the child*. London: Harcourt, Brace, 1932.
4. WALLIS, W. A., & ROBERTS, H. V. *Statistics: a new approach*. Glencoe, Ill.: Free Press, 1956.

THE RELATIONSHIP BETWEEN RIGIDITY-FLEXIBILITY IN CHILDREN AND THEIR PARENTS¹

ABRAHAM BLUM*

Purdue University

The inability of some individuals to alter their mode of response to stimuli with some similarities and some differences has been observed in numerous and varied areas of human activity. It has been suggested that such psychological rigidity² is an important personality factor which circumscribes thinking, limits creativity and problem solving, and predisposes individuals toward theories of racism, prejudice and authoritarianism (1).

Yet little is known concerning the genesis of rigidity. One theory proposes that rigidity is based upon learned behavior patterns and that this form of thinking and activity is transmitted from parent to child through general home atmosphere and parent-child relationships (4).

This study, which was part of a larger one, deals with the assumption that there is a relationship between the degree of rigidity-flexibility displayed by children and their parents. The prediction is that in a comparatively homogeneous group of children relative rigidity-flexibility will be positively related to this same variable in the parents of the group.

METHOD

Sample

The sample consisted of all children 4 years of age or older enrolled in a university laboratory nursery school and of both parents of each of these children. At the outset of testing, this group consisted of 25 children and 50 parents. The final group numbered 17 children and their parents (34

*Department of Child Development and Family Life, Purdue University, Lafayette, Indiana.

¹ Taken from a dissertation presented to the Faculty of the Graduate School of Cornell University (2). A portion of this paper was presented at the American Psychological Association meetings in September, 1954 (3).

² Rigidity is defined in this study as the relative inability to shift set or tolerate ambiguity (as a general characteristic of an individual's behavioral organization). In operational terms, this means the inability to change one's previously developed response patterns when objective conditions warrant such change, the relative inability to perceive and/or respond to a change in a situation, and the relative inability to react in an objectively pertinent manner to tasks which are either ambiguous or choice-forcing.

CHILD DEVELOPMENT

adults). There were 10 girls and 7 boys in this group, ranging in age from 48 months to 63 months. The mean age was 55.6 months, with the mean for the girls, 55.9 months, the mean for the boys, 55.1 months.

In order to avoid as many influences outside the home as possible, it seemed desirable to measure the relative rigidity of still younger children. However, it was also necessary for the children to be capable of at least a minimum of abstraction ability since there is evidence that indication of rigidity at a very young age may actually be a manifestation of inability to "abstract." Therefore, somewhat older children were chosen so that they might meet the proposed abstraction ability criterion; yet, they were young enough that their experience outside of the home was relatively similar.

The 34 adults in this study were all the natural parents of the 17 children described. The mothers ranged in age from 23 to 43 years, with a mean of 34 years. The age range of the fathers was from 25 to 47 years, the mean, 35 years. All the families were from the middle income group; all but two of the fathers were professionals.

Pretest of Abstraction Ability (Child)

To minimize the possibility that indication of rigidity was in actuality a problem of inability to abstract and because of the difficulty of differentiating certain aspects of abstract ability, a pretest technique to measure minimum ability of abstraction was developed.³ The material utilized in this pretest consisted of three paper plates and nine of the 12 Weigl-Goldstein-Scheerer color-form blocks (three red, three blue, and three yellow, each color set having a triangle, a circle, and a square). The instructions suggested by Goldstein and Scheerer (7) for this test were used, with the addition of "(Put those that go together) *together in these plates.*"

S was seated at a nursery school table. Three white paper plates (diameter $6\frac{1}{4}$ inches) were placed in front of S, some six inches from him and three inches away from each other. E placed the nine Weigl-Goldstein-Scheerer blocks on the table between the child and the plates. These blocks were mixed about on the table while E said, "Now they are all mixed; put those that go together, together in these plates" (pointing to each plate).

After S placed the blocks in the plates, E commented, "That's fine." If S did not put all the blocks in the plates but stopped, E asked the child if he was "all done/finished?" If S indicated that he was not, then he was allowed to finish before the next trial was begun. If S did not sort according to color or form, the trial was repeated using the same instructions as before.

If the sorting was based on color or form, E asked, "Why do they go together this way?" A second set of trials was then begun by telling S to "Put them together in a *different* way in these plates." When S sorted the "different way," he was again asked, "Why do they go together this way?"

³ The original Weigl-Goldstein-Scheerer test (7) for abstract-concrete behavior was pretested with a nursery school group similar to the proposed child sample, but was found to be too difficult for them. Therefore, this simplification of the test was used.

ABRAHAM BLUM

S was given five trials for the first sorting (color or form). After completing either one of these sortings, S was given five more trials in order to "shift" to the other type of sorting. If S did not sort by color or form in the first five trials, he received an additional five trials.

The criteria for minimum abstraction ability were: (a) sorting by color or form in one of the first five trials, or (b) where the child did not achieve an initial sorting in the first five trials, he did sort by color or form *and* did verbalize the sorting concept in the second set of five trials. Only those children who met one of these criteria were included in the final study.

Child Rigidity Measure

In order to investigate this problem, it was necessary to have measures of rigidity-flexibility which could be used with both children and adults. Although instruments designed to measure rigidity exist, none was available for use with children below school age. It was therefore necessary to construct such an instrument. Originally an attempt was made to use one instrument with both groups. When this was found to be impossible, two tests were constructed.

Since it has been suggested that young children would best respond to a pictorial test using common, everyday subject matter, a dog-to-cat transition series was devised. The extent to which this test is similar to the instrument suggested by Frenkel-Brunswik (4, 5, 6) is not known.⁴

The Child Transition Test (CTT) was constructed by first drawing the two "end" pictures, that is, a dog and a cat as simple, single line, stylized representations of equivalent size and stature. The transitional pictures were constructed by first superimposing the two original pictures to form the middle picture. This picture was then superimposed on each of the end pictures to form the second and fourth pictures of the series. This procedure resulted in a five card, reversible series, containing a cat, an "almost" cat, a half-cat-half-dog, an "almost" dog, and a dog (Figure 1).

In the test situation each child was asked to identify the figure on each card, presented one at a time. Since the series is considered reversible, the serial order was controlled, so that cat and dog were presented an equal number of times as the first card of a series. After a minimum of seven days had elapsed, each child was retested, the second order being the reverse of the first. This second reverse order was also in reverse color, since photostats were available in both black on white and white on black.

Adult Rigidity Measure

The CTT was not very successful with adults. The relative simplicity of the task gave almost no differentiation during pretesting, and little motivation was apparent in adult subjects. Therefore, with the CTT as a

⁴ In a personal communication Dr. Frenkel-Brunswik indicated that she was not ready to distribute her dog-cat material since her work with these pictures was only exploratory, though definite trends were indicated.

CHILD DEVELOPMENT

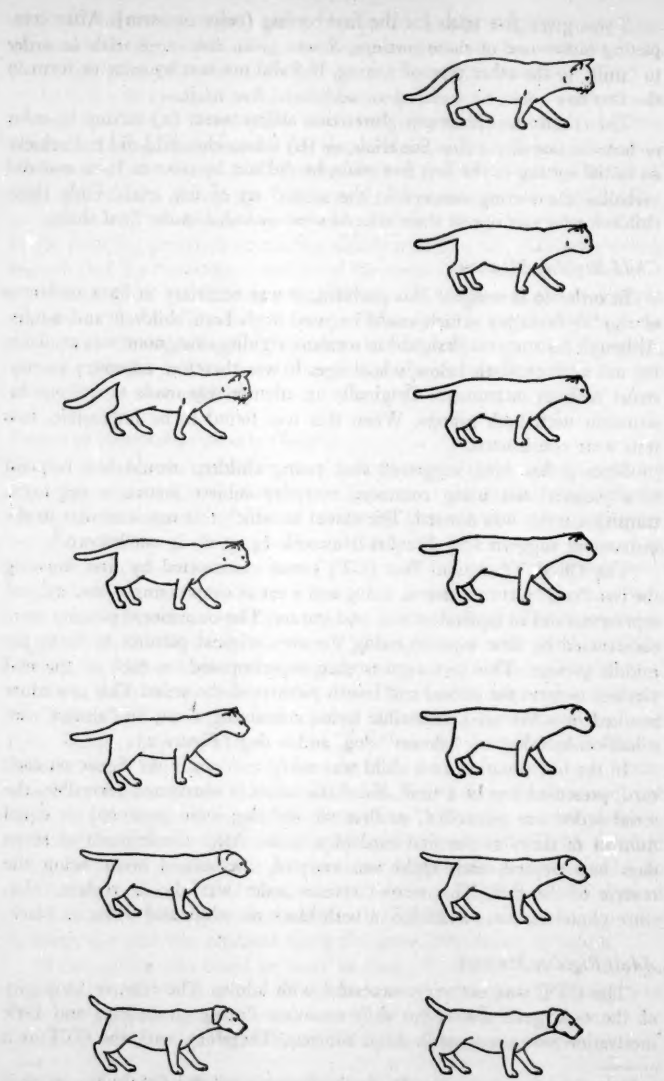


FIGURE 1—The Child Transition Test

FIGURE 2—The Adult Transition Test

base, an adult version was devised by drawing three entirely new transitional pictures in place of the middle card of the CTT. The final Adult Transition Test (ATT) therefore consisted of a seven card dog-to-cat series (Figure 2).

The test procedure was also revised. The adult series was presented in scrambled order, requiring each subject to "order" the cards. After the subject placed the seven cards in an order, he was asked to explain why he had so placed them.

Ranking of Rigidity-Flexibility

Because there was no basis for operationally defining scale units of rigidity-flexibility, the transition series results were used to establish rankings of relative rigidity-flexibility for the children and the adults

Child. The procedure used to establish the children's ranking was as follows: The first ranking was based upon the point of shift, that is, the card at which the child named an animal different from the first animal he named. For example, if a child's responses to the series were "dog, dog, panther, cat, cat," the child would be described as having shifted on card 3. These ranks were then further differentiated on the basis of the total number of shifts made. In the above example, the child had made two shifts, from the dog to panther and from panther to cat. Further breaking of ties within subtests was accomplished by using point of shift and number of shifts in the second testing.

After the final ranking was made, 10 separate ranks were found. (Table 1 gives the results for the points of shift and number of shifts for both tests, for all children.) Six ranks were tied, five with two in each and one with three. (All ties contained at least one child of each sex.) The boys' mean rank was 11, and the girls', 7.5. Using the H test (9) to test for the significance of the difference between the ranks of the boys and girls, it was found that the null hypothesis could not be rejected at an acceptable level of significance.

As can be seen in Table 1, points of shift in the first test covered the entire possible range from one to four (cards 2 to 5).⁵ Six children (35 per cent) shifted on the second card, three (18 per cent) on the third card, six (35 per cent) on the fourth card, and the remaining two (12 per cent) on the last card. All the children shifted at some point in the presentation of the five cards. In the first presentation, there were from one to four shifts during an entire test, again covering the entire possible range. More than half (58 per cent) of the sample shifted three times and 12 per cent shifted four times. It must be noted that the possible number of shifts is directly reciprocal to the point of shift, i.e., if a child shifts at card 5, he cannot have more than one shift, while only a child who shifts at card 2 can possibly have four shifts during the test.

⁵ In pretests with children under 4 it was found that in some cases no shift took place for any of the cards, not even the fifth, which objectively, is a distinguishable dog or cat.

CHILD DEVELOPMENT

TABLE I

CHILD TRANSITION TEST RESULTS AND RANKINGS OF CHILDREN'S RIGIDITY-FLEXIBILITY

(In order of increasing rigidity)

* Child	T E S T 1		T E S T 2		Rigidity- Flexibility Rank
	Point of First Shift	Number of Shifts	Point of First Shift	Number of Shifts	
<i>Girls</i>					
A	4	1	3	1	14
B	2	2	3	1	5.5
C	4	1	2	1	11.5
D	2	4	2	3	1
E	2	2	4	1	4
F	3	1	4	1	7.5
G	5	1	3	1	16.5
H	3	1	3	1	9.5
I	2	4	2	2	2
J	2	3	2	2	3
<i>Boys</i>					
K	3	1	4	1	7.5
L	4	1	3	1	14
M	4	2	2	1	9.5
N	5	1	3	1	16.5
O	4	1	2	1	11.5
P	2	2	3	1	5.5
Q	4	1	3	1	14

The results of the retest show a more limited range. All the children shifted before the fifth card was presented but no child had four shifts in an entire test. As for point of shift, 35 per cent of the children shifted on the second card, 47 per cent on the third, and 18 per cent on the fourth. Fourteen (82 per cent) of these children made only one shift in the second test, while 12 per cent and 6 per cent had two and three shifts, respectively.

In the construction of the CTT it was thought that both the first and second tests might be given equal weighting in establishing a rank order. However, the above results indicate an obvious discrepancy, which the author predicted, based on practice effects from the first test. Therefore, the present method of rank ordering was established before the testing began. These results support such a method as opposed to equal weighting for both tests.

Adult. The ranking of the adults, using the ATT, was based upon two criteria in the ordering of the cards and three criteria in the reasons given for the ordering. In terms of ranking towards the flexible end of the continuum, these criteria were: (a) correct ordering, (b) placement of the seven cards in a single series, (c) verbalization of progression or transition

ABRAHAM BLUM

in the series, (d) consideration of the whole animal versus any specific parts, and (e) absence of expressed confusion or ambiguity on the part of the subject. For example, the adult who placed the cards in correct order, in either direction, and explained that there was a change in the drawings, from a cat to a dog, would be ranked relatively flexible as compared to the adult who made two sets out of the seven cards and explained that he was not sure of his ordering, but it seemed that there were pictures of two dogs and five cats and he could tell this by the difference in the animals' hind legs.

Using the above criteria, four judges and the experimenter ranked the adults' responses. These five rankings were analyzed for agreement by the use of Kendall's *W* (8) for concordance of judgment. *W* was found to equal .802 which is significant beyond the .1 per cent level. The summed ranks from these five rankings were then used to establish the final ranking of the parent subjects. This final ranking, for all 34 subjects, was found to contain 31 separate ranks, with three of these being ties of two individuals each.

RESULTS

Parent-Child Rigidity

The ranking for the mothers on the ATT correlated +.354 (not significant) with the children's ranking on the CTT; the father's ranking on the ATT correlated +.289 (not significant) with the children's ranking on the CTT (Table 2).

TABLE 2
RIGIDITY CORRELATIONS
(Child Transition Test and Adult Transition Test)

Group		Group	r	p
Child	Mother	+.354*	ns
Child	Father	+.289*	ns
Child	Mother and Father	.459**	.05
Mother	Father	-.007*	ns

* Corrections made for ties.

** Coefficient of multiple correlation for three variable problems.

The multiple correlation of mother and father rigidity as measured by the ATT with the results on the CTT was .459, significant beyond the 5 per cent level. This is a particularly interesting finding in light of the fact that the rho for mother and father rankings on the ATT was -.007, which is not significant.

These findings support the prediction, namely, that in a normal sample the child's relative rigidity-flexibility pattern is positively related to this variable in his parents. Although no significant relationships were found for the child and either one of his parents separately, a significant relation-

CHILD DEVELOPMENT

ship was found when both parents were considered. Such findings should further support investigations which attempt to understand and measure the influence and effect of *both* parents, as a family entity, upon the growth and development of the child. The need for such an approach is more apparent when we note that the relationship of relative rigidity patterns between the mothers and fathers in this study was virtually zero.

All this does not, of course, imply that investigation of interrelated family patterns and their effect upon child development will necessarily supply answers as to the actual genesis of such factors as rigidity. Such knowledge can be gained only by study that would provide a test of the assumption that parental rigidity patterns are directly involved in child-rearing practices and that these practices in turn influence the child's rigidity-flexibility pattern. More specifically, such study would concern itself with family authority patterns as the combined outgrowth of both parents' attitudes and activities towards authority and authority figures, social class status and aspirations, ethnic values and prejudices as they may affect the relative rigidity of a child.

In conclusion, it should be stated that the preliminary use of the transition technique indicates that further development and extension of this method promises to provide a useful instrument for measuring psychological rigidity-flexibility.

REFERENCES

1. ADORNO, T. W., FRENKEL-BRUNSWIK, ELSE, LEVINSON, D. J., & SANFORD, N. R. *The authoritarian personality*. New York: Harper, 1950.
2. BLUM, A. H. Parent and child rigidity patterns and their relationship to child rearing practices. Unpublished doctoral dissertation, Cornell Univer., 1953.
3. BLUM, A. H. A technique for measuring rigidity-flexibility in children and adults. *Amer. Psychologist*, 1954, 9, 334. (Abstract)
4. FRENKEL-BRUNSWIK, ELSE. A study of prejudice in children. *Hum. Relat.*, 1948, 1, 293-414.
5. FRENKEL-BRUNSWIK, ELSE. Intolerance of ambiguity as an emotional and perceptual personality variable. *J. Pers.*, 1949, 18, 108-143.
6. FRENKEL-BRUNSWIK, ELSE. Patterns of social and cognitive outlook in children and parents. *Amer. J. Orthopsychiat.*, 1951, 21, 543-558.
7. GOLDSTEIN, K., & SCHEERER, M. Abstract and concrete behavior. *Psychol. Monogr.*, 1941, 53, No. 2.
8. KENDALL, M. G. *Rank correlation methods*. London: Charles Griffin, 1948.
9. KRUSKAL, W. H., & WALLIS, A. W. Use of ranks in one-criterion variance analysis. *J. Amer. stat. Ass.*, 1952, 47, 583-621.



Handwritten signature or text, possibly "George H. ..."



Gertrude L. Rodenbaugh

FLORENCE L. GOODENOUGH

1886 - 1959

Florence L. Goodenough, Professor Emeritus in the Institute of Child Development and Welfare of the University of Minnesota, died April 4, 1959, at the age of seventy-two. In her latter years, despite failing vision, she maintained a courageous spirit and a lively interest in her profession.

Born at Honesdale, Pennsylvania, August 6, 1886, Florence Goodenough early showed her own remarkable intellectual gifts. She was educated in her home to the equivalent of a high school diploma. She received her B.S. degree from Columbia University in 1920 and her M.A. from Columbia in 1921, and served as Director of Research in the Rutherford and Perth Amboy (New Jersey) Public Schools from 1919 to 1921. Stanford University awarded her the Ph.D. in 1924. At Stanford she assisted Professor Lewis M. Terman on the initial phase of his famous studies of gifted children and continued informal collaboration into the later years of these studies. She came to Minnesota in 1924 as Chief Psychologist of the Minneapolis Child Guidance Clinic and in 1925 joined the staff of the newly organized Institute of Child Welfare at the University of Minnesota as Assistant Professor. She became Professor in 1931 and served until her retirement in 1947.

In her own field, Professor Goodenough was one of the most distinguished scholars of her time. She wrote numerous scientific articles exploring many phases of the new and rapidly growing area of child development. She was President of the Society for Research in Child Development in 1946 and 1947 and concurrently was Secretary and then President of the newly formed Division on Childhood and Adolescence of the American Psychological Association. She served in various other capacities in that Association and in the National Council of Women Psychologists. She developed the Goodenough Draw-a-Man Test, which has had world-wide use. Among the many texts she has written, *Developmental Psychology* is

CHILD DEVELOPMENT

most widely known and has been translated into a number of foreign languages. Other books dealt with exceptional children, mental testing, and methods of studying children.

As scientist and scholar, Florence Goodenough was rigorous and exacting; as teacher, counselor, and friend she was human, and humane, though always setting high standards. Devastating in criticism of careless or shoddy work, she was unfailingly patient and helpful wherever she saw promise of growth. Endowed with that wise and balanced perspective which, although exceedingly rare, is sometimes carelessly called "common sense," she enriched all who sought her counsel. She delighted in sharing her wide-ranging interests and avocations with her friends, who were greatly rewarded by her insights. She was an informed lover of birds, music, and flowers, a bibliophile, and an amateur photographer of note.

Her colleagues and students will remember her steady insistence on truth and integrity. They will also recall that, to her, children were not only objects of scientific study—they were also persons of intrinsic worth. To her, parents were not antagonistic to children and to trained professionals—they were fellow seekers after wisdom and after good things for children.

DALE B. HARRIS

University of Minnesota

CHILDREN'S BEHAVIOR UNDER STRESS AND ITS RELATION TO TRAINING BY PARENTS TO RESPOND TO STRESS SITUATIONS¹

EDWIN LEVY*

Bleuler Psychotherapy Center

Stress, frustration, and conflict are important concepts in mental hygiene, and much research has been devoted to aspects of stressful situations and the reactions of individuals exposed to them. This study is in the area of young children's reactions to stress, concerned specifically with the relationship between the children's reactions and certain antecedent conditions: namely, training to respond to stress situations. A natural and realistic stress situation was selected for study, a two-day period of hospitalization for minor surgery. In accordance with principles of learning, children who have been trained for stress in a way judged to be effective should have more adjustive responses in a stress situation than those who have received training judged less effective. The study was constructed to investigate the degree to which children's previous training was related to their behavior in a new and unfamiliar stress situation.

The concepts of stress, frustration, and conflict have often been employed almost interchangeably and distinctions are not clear-cut. Summarizing several points of view, there seems general agreement that these concepts involve (a) behavior motivated towards some goal (whether the goal consists of attainment of an expectation, attainment of material objects, or a

* Bleuler Psychotherapy Center, 139-24 88 Avenue, Jamaica, New York.

¹ This study is based on a doctoral dissertation submitted by the author to the Graduate School of Adelphi College. The author wishes to thank Dr. Gordon Derner, Dr. Robert Cranston, and Dr. Gerald Lesser for their help.

CHILD DEVELOPMENT

state which might be thought of as psychological equilibrium); (b) blocking of the goal-directed activity by some barrier; (c) continuation of the drive towards the goal; (d) increased tension arising from blocked activity with the possible arousal of emotional behavior (2, 3, 6, 10, 18).

In stress experimentation attention has been directed towards the problem of having the stress a realistic one (10). Predictability of subjects' reactions has also been investigated with Dollard, Doob, Miller, Mowrer, and Sears (7) interpreting postfrustration behavior in terms of direct aggression, displaced aggression, or substitute behavior. Later, Miller, Sears, Mowrer, Doob, and Dollard (16) suggested a revision of the original premise to read, "Frustration produces instigations to a number of different types of response, one of which is an instigation to some form of aggression" (p. 338). Direction of reactions according to the experimental observations of Barker, Dembo, and Lewin (1) is seen in terms of regression. This position is commented upon by Child and Waterhouse (3, 4) who put forth evidence for an interference hypothesis, rather than supporting the concept of regression.

Maier (13) sees postfrustration behavior as nonmotivated behavior without a goal, a position developed from his experiments with rats forced to respond in insoluble problem situations. Davitz (6), however, finds that Maier's basic data may be analyzed in terms of motivation and learning theory, and states that it is not necessary to raise a new theoretical issue of nonmotivated behavior. The general conclusion which may be reached is that there is no over-all factor in frustration or stress per se which determines the direction of behavioral change, and no single direction which may be anticipated. The effect that stress has on behavior seems multidetermined and reactions in turn would appear multidirectional.

In previous research, the relation of current behavior to the past history of the individual is one aspect yet to be explored. Investigators have seen this aspect of stress experimentation as a potentially fruitful source of research. Lazarus (10) recommends as part of future exploration a study of "... the relationship between personality variables, past history and performance under stress" (p. 315). Child (4) refers to the importance of past history in a statement that reactions to a frustration "... seems likely to be influenced by what particular response tendencies to this changed situation have become dominant as a result of ... previous experience in similar situations" (p. 134). Sears (19) raises as a research question "... the determination of the specific factors which cause one kind of frustration-reaction rather than another to occur" (p. 345).

Child (4) also comments that the habits of responding to stress and frustration can have a bearing upon the individual's reactions. He reports that training to respond to stress is one of the assumptions of military training, so that a higher quality of performance may appear under stress, rather than a lower. "[There is] ... the assumption that training can modify the way a person responds to an intense emotional state ... so ... that intense

EDWIN LEVY

emotion may come to have an organizing rather than a disorganizing effect on behavior" (p. 127).

The question of etiological factors in stress reactions may be conceived of as a problem in experience and learning. Directly pertinent is a study by Hymovitch (9) on problem-solving in rats. A group of laboratory rats was reared in two ways: one group in a "psychologically restricted" environment, consisting of a small cage with food, water and an opportunity for exercise, but with no problems to solve, no pain and no need to get along with other rats; the second group in a "free" environment, consisting of a large box with obstacles to pass, blind alleys to avoid, other rats to get on with, and with great opportunity for problem-solving and great need for learning during growth. At maturity the problem-solving ability of the animals was tested with the conclusion that the rats brought up in the shielded environment had a lasting inferiority of skill in problem-solving. The study suggests a relationship between prior exposure to problems and behavior in a problematic situation.

An investigation specifically designed to examine the results of previous training on postfrustration behavior in young children was conducted by Davitz (6). He studied the differential effects of aggressive training and constructive training on the responses to frustration made by children 7 to 9 years of age. It was hypothesized that the training received would develop in each group a specific behavior tendency, and that the behavior tendency would be related to that physical setting. It was also hypothesized that the learned behavior tendencies would differentially affect the behavior of the two groups following frustration.

The experimental population consisted of 40 children selected from a group of children in residence at a summer camp. All children received seven 30-minute training sessions in a playroom, with half the total group receiving training designed to encourage aggressive reactions and the other half receiving training designed to encourage constructive reactions. Frustration was introduced experimentally by interrupting a film being watched by the children in the playroom and taking away candy bars which they had just been given. Analysis of subsequent behavior supported the conclusion that "... previous training in situations similar to that in which frustration is encountered is a significant determinant of the organism's post frustration behavior" (6, p. 314).

The study by Davitz appears to be the first instance of systematic research on the problem of training and behavior under stress in childhood. It shows the feasibility of investigating in experimental situations factors from past history which may have bearing upon present behavior.

PURPOSE OF THE PRESENT STUDY

The present study is specifically concerned with the effectiveness of training for stress experiences in childhood. The question investigated is

CHILD DEVELOPMENT

whether a general kind of training can take place in learning to respond to stress. It seems likely that children trained to react in one setting will show positive transfer of principles and hence respond adequately in another, albeit unfamiliar setting.

Miller and Dollard (15) state that effective methods of eliciting the first correct responses from organisms so that they can be rewarded include the conditioning technique, imitation, and verbal instruction. This study is primarily concerned with the verbal teaching by parents to respond to stress, and with the effectiveness of this aspect of child-rearing. Where parents have provided responses to cues, and these responses have been judged to be effective by independent judges, the child's behavior during stress should be adjustive to the situation. Where the effective responses to stress cues have not been taught, the child's behavior under an unfamiliar stress might be expected to be maladjustive to the situation. Adjustive behavior is defined as that which indicates the child is accepting of the situation although under some tension and capable of carrying on social participation. Maladjustive behavior is defined as that which is characterized by much trial and error, unhappiness, and overt emotional disturbance.

METHOD

The present study measures past training to respond to stress by means of controlled interviews with parents and measures the effects of this past training by means of controlled observations of children. A situation was sought which would represent a realistic stress in childhood, and hospitalization was chosen since it seemed to contain the elements which experimentation has shown to be primary in producing stress (5, 17). The elements present are first, that the situation is stressful when it is not under the subject's control, and second, that stress which must be dealt with immediately is perceived as greater than that which can be postponed. Almost all hospital situations in childhood contain elements of stress according to these criteria. The child is essentially helpless to control his environment and the situation is one that must be dealt with immediately. Further, the child is usually in strange surroundings and apart from his family.

Maslow (14) comments that "threat includes phenomena that are subsumed neither under the head of conflict nor frustration as these words have been used in the past" and cites illness in young children as an example (p. 83). The fact that the sense of stress in a situation may be heightened by the organism's helplessness was observed by Mowrer. Using rats as subjects in an experimental situation he found greater fear-reactions occurring in those rats which could not experimentally terminate a shock than in rats which could. He comments that "... fear aroused by a physical pain is a function of whether the pain is or is not under the subject's control" (17, p. 200). Combs and Taylor state that the sense of threat may be increased by the physical and psychological immediacy of the threatening

situation and comment "... it would appear that a threatening perception which must be dealt with at once is more threatening than one which could be postponed" (5, p. 421).

In experimental design on the preparation for a hospital stay and children's reactions once in the hospital, concentrating on a single kind of illness with its concomitant medical and surgical procedures permits for increased controls. The children's stay in the stressful situation becomes more standard and care and other factors become more uniform. The situation finally selected for study was hospitalization for tonsillectomy and the study was carried out in a large children's hospital.

While on the Tonsillectomy and Adenoidectomy Service all children undergo the same routine with minor modifications. They are seen by the same group of physicians when recommended for operation and go through the same medical routine. The period of hospitalization is the same for all, lasting an afternoon, a day, and the following morning. The children used in this research were between the ages of 4 years and 8 years, 9 months. There were 39 children in all, 16 boys and 23 girls. The modal age for both boys and girls was between 6 and 7 years, and there were more girls than boys at both the 7- and 8-year levels. Very few families with children under 4 years used the T. and A. Surgical service during the time the study was conducted, and for varying reasons those that did so did not meet the requirements for control so that none was included in the sample.

Because the use of the T. and A. Clinic Service and the ward facilities is limited to families of low income, the economic level of the population fell within a narrow range. The Service handles both Negro and white children and both were considered in the study. There were 11 Negro boys, 4 white boys, 23 Negro girls and 1 white girl. To insure that the hospitalization of concern in the study would represent a new and unfamiliar stress, children with previous overnight hospital experience were not included in the sample population. Whether the children came from families with one or more siblings, or whether they were only children was also considered. In the investigations prior to the main study it was found that in the families of the economic level of the study the majority of the children had siblings and it was decided to omit those children who did not. Therefore, all children in the present study had siblings. Judging from clinical observation, none of the children showed gross manifest emotional disturbance or gross intellectual retardation.

The method of data collection was evolved to obtain information first on parent behavior in regard to training for general stress situations and to training and preparation for the hospitalization. Second, the children's reactions during the new and unfamiliar hospitalization were observed. All of the parents of the experimental population were contacted and interviewed during the children's two-day stay at the hospital. At the initial contact with parents on the day of admission, it was explained to them that the hospital was conducting a survey on children's experiences around

CHILD DEVELOPMENT

hospitalization and medical care and that their participation as parents was earnestly requested. Most parents agreed to discuss their children and their children's experiences, and at that point an appointment was made for the parent to return for an interview.

The observations on the children during their hospitalization were carried out by the nursing staff of the ward used by the Tonsillectomy Service. The nursing staff was extremely cooperative during the data collecting periods, and since the charge nurses did not change during this time, there was continuity in the personnel most involved in obtaining data. Similarly, at the point where observations on the children were made outside the ward (when the children were in the recovery room following operation) one nurse made the observations on all the children since she was always on duty when they returned from the operating room.

Parent Measures of Training for Stress

To investigate the kind of past training children had experienced with respect to stressful situations, an estimate of parent behavior was essential. A controlled interview was developed and pretested through a series of trial interviews, and the interview schedule was evolved to cover areas of parent behavior which seemed pertinent to training in stress. It was composed of questions designed to elicit from parents the nature of their behavior in regard to general training for stress experiences and training for the specific stress of hospitalization. An attempt was made to frame questions in such a way as to avoid giving the idea that any type of behavior was favored.

The interview was designed to cover an estimate of the child's experience with general stressful situations in the past and training by the mother or others to respond, an estimate of the training to respond to the specific stress situation of hospitalization for tonsillectomy, and an estimate of the child's present adjustment and functioning. In the trial interviews prior to the main study it was found that all areas could be covered in the conversation in a natural way although not necessarily in the order presented in the scale.

The questions in the section on general training for stress were designed to elicit information about the child's experiences with stressful situations and the kind of training provided. The categories covered experiences such as would occur to most children in the course of everyday living and to which they learn to adjust. Included were experiences of being left with adults besides the parents or having baby sitters, being left overnight and separated from the mother, getting accustomed to the dark, to animals, and to strangers. Also discussed were the child's experiences of adjusting to the dentist, coming for medical examinations, beginning kindergarten or nursery school, going to first grade, going alone to the store, and riding alone on the bus. Questions about parent behavior were aimed at exploring what the parent did in relation to the child and the situation. For example, when

EDWIN LEVY

inquiring about the child's experiences of being left with other adults, such questions as these were asked: "Has your child had the experience of being left with someone besides yourself? Who? How much do you think he knew about this ahead of time? What did he do? Who told him? What did you do? How did he respond?"

The questions in the section on preparation for hospitalization attempted to elicit the nature of training about the specific areas to be experienced. This section covered what was done in the way of imparting knowledge and preparation about various aspects of the hospital experience. For example, the parent could tell the child that he must remain in the hospital overnight, and that the total hospitalization required two nights; that his tonsils would be taken out; that he would be put to sleep although the specific procedure of anesthesia need not be outlined; that his throat would hurt after the operation was over but there would be no pain during the operation; that he would be given ice chips after the operation, not liquids or ice cream; and that he could be visited by his parents or other adults. The child's knowledge of the reason for hospitalization was explored through such questions as: "How much do you think your child knew about why he came here? How did he learn it? From whom did he learn it? What did you do or tell him? Do you think he knew he was coming in to have his tonsils out?" In inquiring about specific areas such as the child's preparation for remaining overnight, such questions as these were asked: "Do you think he knew he was going to stay overnight? How did he learn about it? What did you tell him about it?"

The section on present adjustment included a brief developmental history, originally considered useful in starting the interview and in learning who cared for the child in infancy and how this was done. A survey of behavior covered feeding, dressing and toilet habits to learn the nature of the child's dependency in these areas upon others, social adjustment with other children as an indication of his ability to get along with peers, the presence of fears, mannerisms or behavior otherwise symptomatic of anxiety, and school adjustment to estimate the child's relationship to the authority represented by the teacher and to children in a structured situation. The information covered in this section on present adjustment was not essential to the hypotheses but was obtained to permit an evaluation of the child's behavior on follow-up visits to the hospital.

A total of six judges took part in the analysis of the data of the parent interviews and the observations on the children. All of the judges were trained clinical psychologists experienced in working with young children in either a child guidance or hospital setting. Two of the judges were psychologists at another children's hospital in Philadelphia, two were psychologists at a Philadelphia child guidance clinic, and two were on the staff of the children's hospital where the study was done. All of the psychologists had either attained the doctorate or were close to completing the requirements for it.

CHILD DEVELOPMENT

The information obtained from the parents was evaluated by the judges according to the rating criteria presented below. Later it was categorized as to whether it represented Effective or Less-Effective training. Effective training was defined as training by a person who was assumed to be familiar with the situation. It included parents and other adults, and in some instances children where the child had first-hand experience in the situation. Training was considered Effective if:

- a. the parent took the initiative in helping the child understand a situation and respond to it.
(*Example*—"When I saw he had to come to the Clinic I told him we were coming and explained what the doctor was going to do."
"I said he would be put to sleep and wouldn't feel anything when his tonsils came out.")
- b. the parent was relaxed in helping the child understand and respond to a situation; provided responses and explanations when the child asked.
(*Example*—"When I took him to the dentist he wanted to know what the dentist would do and what would happen. I explained about the drill and how the dentist would look in his mouth.")
- c. another adult who is familiar with the situation and could be expected to help the child understand it did so.
(*Example*—"Before he left kindergarten the teacher told them all about the 1st grade and how they would have home work and what it would be like." "His uncle had his tonsils out very recently and told him the tonsils come out, that's what they do in the hospital.")
- d. another child who is familiar with the situation and could be expected to help the child understand it did so.
(*Example*—"He learned how to go to the store because his sister took him and she showed him what to do." "His little friend had his tonsils out and told him he had to stay overnight.")
- e. the situation might be considered under the heading of "positive laissez-faire." For example, the situation might be the kind in which no action by the parent was required because the child was already adjusting well. Or the parent taught the child nonverbally by providing an example the child could imitate.
(*Example*—"He's always slept in the dark. I didn't have to do anything special to get him used to it because he was never scared at all."
"He's been to the clinic so much that he's seen the doctors and nurses often. He already knew that.")

Less-Effective training was defined as the absence of training or training by persons the judges considered likely to be misinformed themselves. Training was considered Less-Effective if:

- a. the situation might be considered under the heading of "negative laissez-faire." For example, the situation might be the kind in which some change in the child's adjustment was needed but the parents did nothing to bring it about. Or the child learned largely through trial and error.

EDWIN LEVY

Or the parent did not know how the child learned to respond to a situation.

(*Example*—"I don't know how she learned to go to the store. I guess she just figured it out by herself." "I don't think he knew anything about staying two nights. I didn't know myself so I couldn't tell him before.")

- b. the parent gave false or incorrect information to the child or misled the child. This may have been either deliberate or through misinformation on the part of the parent.

(*Example*—"If I know he's going to get a needle at the hospital I don't tell him at all. I tell him it won't hurt. I don't want to frighten him.")

- c. another adult gave false or incorrect information to the child, either deliberately or through misinformation.

(*Example*—"His aunt told him about when she had *her* teeth fixed. She said they always give gas so it don't hurt. She didn't know they do it different for children.")

- d. another child gave false or incorrect information to the child either deliberately or through misinformation.

(*Example*—"The little boy down the street said there was a boogie man at night. He likes to tease. I don't know why children do those things to scare each other." "Her little friend told her the tonsils were frozen out. And that she could keep them afterwards.")

Interjudge agreement was computed on the interview information about general training for stress experiences, and on the specific training for hospitalization. Agreement was evaluated according to whether judges rated a reported incident of parent-behavior as falling into the category of Effective or Less-Effective training. Agreement was determined using tetrachoric r . Interjudge agreement for the total of 429 items referring to general training ranged from .85 to .99, values which are significant at the 1 per cent level. For the total of 273 items referring to specific training for hospitalization, interjudge agreement ranged from .89 to .95, values which are significant at the 1 per cent level. There was consistently high agreement as to whether the incidents reported by the parents could be considered examples of Effective or Less-Effective training.

Criteria Measures of Behavior during the Stress of Hospitalization

Experimental work leading to theoretical formulations of the reactions to threat and barrier situations has been done by Lewin (12) and Rosenzweig (18). One application of the theoretical framework to a hospital situation was done by Leitch and Escalona (11). Their observations on the behavior of infants under the stress of hospitalization suggest that practical use may be made of the theory in this situation.

Lewin describes reactions to tension in a barrier situation in terms of going-out-of-the-field and seeking escape (which may occur through fantasy or play), taking action against the barrier (which may be seen in strife

CHILD DEVELOPMENT

with the adults establishing the barrier, or strife against the barrier, or strife against a task), encysting and defiance (where the child may withdraw into himself, or show fighting, defiant attitudes), and emotional outbursts (which may manifest themselves in diffuse emotional discharges).

Rosenzweig postulates the direction of reactions in a tension situation as extrapunitive (in which an individual attributes a frustration to others and may show anger and resentment), intropunitive (in which an individual attributes the frustration to himself), and impunitive (in which there is complete avoidance of attaching blame anywhere). In a study dealing entirely with overt reactions of hospitalized children on an observational basis, Winkley (20) described the children's reactions as positive (implying that the child seemed pleased by the situation or really accepting of it), negative-aggressive (implying that the child seemed hostile, retaliatory, resentful or rebellious), and negative-passive (implying that the child seemed fearful, withdrawn or clinging).

Considering these various categorizing systems, it seemed possible to conceive of children's behavioral reactions to a barrier and stress situation within the active-passive and positive-negative limits described by Winkley, and with the underlying theoretical meanings of Lewin and Rosenzweig. For example, behavior which was both positive and active might be observed in situations in which a positive valence had been induced (Lewin), and where reactions might be described as impunitive since there would be no appearance of attaching blame (Rosenzweig). Overtly the child would appear pleased or really accepting of the situation (Winkley).

Behavior which was both positive and passive would be similar to the above, but might be thought of as occurring in a situation in which the child needed more help and encouragement and in which he would respond to the security of directions.

Behavior which was negative and active might show action against barriers with strife with adults, tasks or barriers (Lewin), defiant fighting attitudes with the child seeking to leave the field bodily (Lewin), or extrapunitive actions with the child angry, or resentful of others (Rosenzweig). Overtly the child might appear noncooperative, hostile, retaliatory, resentful or rebellious (Winkley).

Behavior which was negative and passive might be thought of as encysting, in which the child withdraws into himself (Lewin), or as intropunitive, attributing the frustration to himself (Rosenzweig). Overtly this might appear as fearful, withdrawn, or clinging behavior (Winkley).

Behavior which showed diffuse emotional discharge might be characterized by explosive crying with no real attempt at action (Lewin). The child might be assumed to be overwhelmed by the situation and to have temporarily given up the struggle.

The positive or negative quality of a child in a new stress situation was assumed to be affected by the training and preparation to enter the situation. Training judged to be effective was assumed to bring about posi-

tive reactions since the child had responses available. This might be seen in a readiness to accede to required behavior or in effective social participation. Training judged less effective was assumed to result in fewer adjustive responses and in more negative, trial-and-error, or disorganized ones.

Transferring these concepts into the specific behavior to be used as criteria required a knowledge of the hospital situation. Several weeks were spent in observing the complete procedure followed on the tonsillectomy service: discussing with the physicians, nurses, and other personnel directly involved with the children and observing the children themselves as they were admitted to the hospital, were separated from their parents, underwent anesthesia, were sent to the recovery room following operation, etc., throughout the two-day period. These observations led to a timetable of activities which was the same for each child with only minor variations. From this timetable a series of representative stress points was selected which allowed for observation in a number of situations at frequent intervals during the day. The stress points selected for the first day included the initial reaction to the ward situation (from 10:00 A.M. to 12:00 noon), rest hour (1:00 to 2:00 P.M.), play time (2:00 to 4:00 P.M.), supper (4:30 to 5:30 P.M.), bed time (7:30 to 8:00 P.M.) and sleep (10:00 P.M. to midnight). For the second day the stress point selected for observation included the morning medications (6:30 to 7:00 A.M.), reactions postoperatively in the recovery room (10:30 A.M. to noon), reactions on the ward following operation (2:00 to 3:00 P.M.), bed time (7:30 to 8:00 P.M.) and sleep (10:00 P.M. to midnight). For the third day, the reactions observed were at breakfast (7:00 A.M.) and the reactions to the mother's arrival and going home (any time from 9:00 A.M. to noon).

For each of these stress points five descriptions of behavior were constructed, intended to represent the categories of Positive-Active, Positive-Passive, Negative-Active, Negative-Passive and Diffuse Emotional Discharge. For example, around the first stress point (Initial Reaction to Ward Situation) the following behavior descriptions were constructed:

Positive-Active—Friendly with nurses and children. No difficulty in leaving mother or "settling down" on ward. Makes spontaneous overtures or comments to others.

Negative-Active—Noisy, angry. May not want to be left; or may not want mother to leave. May demand to be taken home; or cry noisily.

Positive-Passive—May cry gently but can be soothed by mother or nurse. With help and friendliness from others becomes quiet.

Negative-Passive—Quiet but frightened and apprehensive. May remain close to mother or nurse. May remain by self without making overtures to others.

Diffuse Emotional Discharge—Noisy, prolonged crying. May pay no attention to others. May not respond to overtures of others. Difficult to soothe.

A total of 65 items were constructed, five for each of the 13 stress points. These descriptions of overt behavior were next submitted to judges who

CHILD DEVELOPMENT

were asked to match the constructed behavior descriptions with the five categories derived from theory. For the total of 65 items interjudge agreement was computed using the coefficient of contingency. The maximum value of C in a 5×5 table is .89, and interjudge agreement ranged from .87 to .88, values which are significant at the 1 per cent level.

The final step in establishing the criteria consisted of drawing up the check list of behavior which covered each stress point during hospitalization. For example, the check list for the Initial Reaction to the ward situation was as follows:

Friendly	Leaves mother with difficulty	Resists undressing
Crying	Leaves mother easily	Changes clothes easily
Noisy	Demands to go home	Protests but then changes clothes
Angry	Not want mother to leave	Remains close to mother
Crying	Remains close to nurse	Remains by self
Crying noisily	Pays no attention to others	Resists soothing
Frightened	Responds to soothing	Quiet
Prolonged crying	Friendly with children	Friendly with nurse

When the check lists of actual behavior observed during the children's hospitalization were obtained, the information was given to four of the judges. The judges were asked to match each observation of overt behavior at a given stress point with one of the behavior descriptions which had been constructed for that stress point. The matchings of each judge were then compared to note whether there was agreement on the assignment of a specific item in the direction of Positive behavior or Negative behavior. The results of interjudge agreement were measured using tetrachoric r . For the total of 507 items the judges reached agreement ranging from .92 to .95, all significant at the 1 per cent level.

In summary, the method was evolved to obtain information on parent behavior in regard to training for general stress experiences, and to training for the stress of hospitalization. The children's behavior at several crucial times during the new and unfamiliar stress of hospitalization was observed. It was then possible to relate parental training to respond to stress and the children's overt behavior during hospitalization itself.

RESULTS AND DISCUSSION

The basic hypotheses of the study were tested by evaluating the relationships between the interview data and the children's behavior during hospitalization. Hypothesis 1 stated that children who have training judged to be effective in responses to a variety of stressful situations will show more adjustive responses in the unfamiliar stress of hospitalization than those children who have received training judged less effective. To find

a child's position in the group in regard to his general training for stress the following procedure was employed. A standard of agreement by at least three out of the four judges was required on each item for placing the training the child had experienced in either the Effective or Less-Effective direction. Items which did not reach this level of agreement were omitted from the final tabulations. Items on which at least three out of four judges agreed as indicating Effective training were given a weight of 1. Items on which at least three out of four judges agreed as indicating Less-Effective training received a 0 score. A child's position in the group was determined by adding up his scores on the 11 items of general training which could run from 0 (indicating training judged consistently Less-Effective) to 11 (indicating training judged consistently Effective.)

Hypothesis 2 stated that children who have received training judged Effective in responses to the specific stressful situation of hospitalization will show more adjustive responses in that stress situation than those children who have not received such training. A similar standard of inter-judge agreement was required as on the items of general stress training. A child's position in the group in regard to being trained for hospitalization was determined by adding up his scores on the items which could run from 0 (indicating training consistently judged Less-Effective) to 7 (indicating training consistently judged Effective).

The decision as to whether an item of behavior in the hospital indicated Positive or Negative reactions was made on the following way. A standard of agreement by at least three out of four judges as to the categorization of the behavior at each stress point (such as admission, play time, bed time, recovery room, etc.) in the Positive direction (Positive-Active, Positive-Passive) or Negative direction (Negative-Active, Negative-Passive, Diffuse Emotional Discharge) was used, or the item was omitted from further consideration. Items which at least three out of four judges agreed as indicating a Positive reaction were given a weight of 1. Items on which at least three out of four judges agreed as indicating a Negative reaction received a 0 score. A child's position in the group was determined by adding up his scores for each stress point in the hospital which could run from 0 (indicating consistently Negative reactions in the hospital) to 13 (which could indicate consistently Positive reactions in the hospital).

The results indicated that the training and preparation children receive for general stress situations appears to bear little relationship to their behavior during a particular stress situation such as hospitalization; using the median test a p value of .55 was obtained. When training is more specific and related to one experience such as hospitalization, the relationship between that training and subsequent behavior is higher, as shown by the p value of .05.

These results give support for a transfer theory of specific elements. However, in terms of the plan of the study it should be kept in mind that the general training items covered many different situations and possibilities

CHILD DEVELOPMENT

for stress whereas the specific training items were concerned only with the training for hospitalization. Since the information obtained from the parents covered specific aspects of how they had prepared their children for hospitalization and the observations on the children took place around these same aspects during the experience, the possibility of observing transfer effects was enhanced.

It might be said that relevant stimulus-response training is an effective form of verbal pretraining prior to an experience. In effect, the children were helped to attend to the distinctive characteristics of the new situation, perhaps aided in their attending by heightened anxiety. What the parents did in the way of teaching and preparing was accompanied by interest on the part of the children in seeking appropriate and anxiety-reducing responses. The small relationship between general training for stress and behavior during hospitalization yields little support for a theory of transfer of principles or performance-set. However, because of the complexity of the situation studied, one might want to take into consideration other variables such as personality organization which were beyond the scope of the present investigation.

Results indicated a correlation between the training parents provide for general stress situations and the training provides for hospitalization, as shown by a C of .38 and a p value of .01. The correlation is a significant one, although somewhat low, so that one could not predict from parental practices and behavior in general situations which parents would also train their children for a specific situation.

Regardless of the preparation for hospitalization or the absence of it, most of the children appeared to go through the entire experience without marked distress. In general, their behavior during hospitalization appeared to follow a pattern. On the day of admission following separation almost all of the children were observed as quiet, conforming, playing and apparently accepting of the situation. The first upsurge of any kind of tension appeared at sleeptime when some children were observed to be wakeful and to require help and encouragement from the nurses. All of the children were able to fall asleep, however, and no child was sufficiently upset to require sedation.

The greatest range of behavior reactions occurred on the second day when the children experienced the most extreme stress; they were given medications, underwent anesthesia, and felt pain following the operation. The most dramatic upsets occurred in the recovery room where the children were sent immediately following operation prior to returning to their beds on the ward. Upon being awakened in the recovery room, about half of the children remained quietly on the litter, were attentive and responsive to the nurse, and frequently returned to sleep after some slight conversation. The other half showed signs of extreme agitation and distress, screaming and crying and attempting to crawl off the litter, not heeding the nurse, and needing to be physically restrained. The reactions at this time were examined

EDWIN LEVY

for possible relationship to the child's age, sex, prior adjustment, or specific preparation for hospitalization. None of these factors showed a significant statistical relationship to the reactions in the recovery room so that they were apparently mediated by factors outside the scope of this study. The reactions in the recovery room also did not bear a significant relationship to the child's behavior during the total hospitalization. Finally, on the third day of hospitalization, the children as a group were much like the first: quiet, conforming, and becoming playful although in a subdued manner.

The finding that most of the children appeared to go through the hospitalization experience without signs of marked distress has been discussed in the results of other studies on the particular stress of hospitalization and raises several possibilities. One possibility is that children are able to conceal evidences of distress while undergoing certain kinds of stressful experiences but reveal them at a later time when they are once again in the family situation. This is a possibility upon which some light might have been thrown had it been possible to conduct an effective follow-up on the children. This could not be done systematically because many of the parents did not keep their appointments for the follow-up clinics. The few mothers from whom information was collected postoperatively almost uniformly reported clinging and subdued behavior for about a week following the operation with gradual resumption of normal interactions and play patterns. The children did not talk of their experiences to their parents and none of their parents observed their solitary or group play in which the children might have expressed any feelings about the stressful experience they had undergone. One area which might be considered in future research is the extent to which children react to and assimilate a stressful experience after the actual happening has passed. Perhaps, the manner of investigation might be through play, following the observations of Anna Freud and Dorothy Burlingham (8) who noted that children undergoing bombings in wartime England did not always show fear during the experience but afterwards repetitively played out what had occurred.

Another possibility concerning the relatively few indications of distress during hospitalization is that the children were actually able to deal with the situation as it arose. This may have been particularly true for the population from which the children for the present study and other similar studies on the stress of hospitalization were drawn. In this predominantly lower-economic group both parents and children seemed to accept unpleasantness and discomfort as a natural part of daily living. This was evident from the conversation and observations of nurses on the private and semi-private services as compared with the ward service where this study was conducted. The impression of the nurses was that both parents and children on the private services seemed more alert to the distresses and dangers of any hospitalization than those on the ward service. The parents appeared slightly more anxious and solicitous and the children tended to react more overtly and violently to pain than did the ward populations. However, it

CHILD DEVELOPMENT

would be wise to subject the nurse's impressions and conjectures to more rigorous scrutiny before accepting them.

Another possibility is that Negro children in a hospital staffed predominantly by white physicians and nurses may conceal their feelings for fear of disapproval. Many of the Negro parents in their initial visits to hospital clinics give the appearance of being silent, uncommunicative and unfeeling, whereas they may actually be extremely uneasy, concealing feelings and anxieties which appear when a closer relationship is formed with the staff. The pattern of reserve in feelings is one which might be communicated easily to children who would then hide their own distress in a hospital situation.

Concerning the specific preparation and training by parents, the parents of this present research group gave information in their interviews which indicated that few of them prepared their children for stressful experiences in the way anticipated by the experimenter. The majority of the parents perceived stressful experiences as part of living without seeing them as unusually traumatic or feeling that the children needed special preparation or training. There was less parent-child interaction by way of verbalizing, answering questions and explaining, and apparently more by way of setting emotional tone for undergoing an experience. This appeared to be true for families in which parent-child relations seemed poor as well as those in which parents seemed warmly loving and supportive. As one mother of a 7½-year-old girl remarked when asked about how much her child knew about medical examinations before she came to the hospital, "I just bring her in for them. I don't have to tell her anything. . . . I don't make a habit of doing all that explaining. I figure *I'm* the mother and *she's* the child and she does what she's told." This was a mother who, despite her attitude of nonpreparation, had told her child all about the experience of hospitalization in a way which would be considered very acceptable according to the most verbal standards. She simply had not thought of her conversations with the child as "explanations." This child's reactions, incidentally, fell within the upper range of good adjustment, according to the standards of the present study.

It was apparent that there were variations in the motivation of parents to have their children undergo surgery. In many instances it was the school that had diagnosed infected tonsils and urged remediation. The school thus provided the major impetus for the parents to arrange for the operation. The families themselves frequently had not taken action spontaneously and the selection of children growing out of the school's recommendation had an important effect on the age range of the population which was finally available. Whereas in the private and semiprivate services the children were predominantly in the preschool age range the ward group was significantly older. Since many children in the sample population had lived with their inflamed or infected tonsils for a long period, sometimes for years, there was no rush on the part of the parents to secure surgical treatment, and

EDWIN LEVY

hospitalization was undertaken in a leisurely manner. Thus, although the children underwent stress during their separation from home and during the hospitalization, they were less encumbered by the anxieties and fears that beset parents who take younger children for tonsillectomies.

CONCLUSIONS

The study suggests that when parents train their children to respond to general stress situations there is no marked or significant effect on the child's reactions to a specific incident or stress situation such as hospitalization. However, when parents train children to react to a specific stress such as hospitalization, the reactions are more favorable, and the special preparation and training seem to make the children more comfortable in that situation.

The study also suggests that most children take the stress of hospitalization for tonsillectomy in their stride with few reacting poorly. The exception occurs postoperatively in the recovery room where the greatest number of disturbed reactions were observed. At other times during hospitalization the children seemed to come through without indications of being upset.

REFERENCES

1. BARKER, R. G., DEMBO, TAMARA, & LEWIN, K. Frustration and regression. In R. G. Barker, J. S. Kounin, and H. F. Wright (Eds.), *Child behavior and development*. New York: McGraw-Hill, 1943. Pp. 441-458.
2. BRITT, H. S., & JANUS, S. Q. Criteria of frustration. *Psychol. Rev.*, 1940, 47, 451-469.
3. CHILD, I. L., & WATERHOUSE, I. K. Frustration and the quality of performance: I. A critique of the Barker, Dembo and Lewin experiment. *Psychol. Rev.*, 1952, 59, 356-362.
4. CHILD, I. L., & WATERHOUSE, I. K. Frustration and the quality of performance: II. A theoretical statement. *Psychol. Rev.*, 1953, 60, 127-139.
5. COMBS, A. W., & TAYLOR, C. The effect of the perception of mild degrees of threat on performance. *J. abnorm. soc. Psychol.*, 1952, 47, 420-424.
6. DAVITZ, J. R. The effects of previous training on postfrustration behavior. *J. abnorm. soc. Psychol.*, 1952, 47, 309-315.
7. DOLLARD, J., DOOB, L. W., MILLER, N. E., MOWRER, O. H., & SEARS, R. R. *Frustration and aggression*. New Haven: Yale Univer. Press, 1939.
8. FREUD, ANNA, & BURLINGHAM, DOROTHY. *War and children*. New York: War Medical Books, 1943.

CHILD DEVELOPMENT

9. HYMOVITCH, B. Problem-solving in the rat. *J. comp. physiol. Psychol.*, 1952, 45, 313-321.
10. LAZARUS, R. S., DEESE, J., & OSLER, SONIA F. The effects of psychological stress on performance. *Psychol. Bull.*, 1952, 49, 293-316.
11. LEITCH, MARY, & ESCALONA, SIBYLLE. The reactions of infants to stress: a report on clinical observations. In Anna Freud, H. Hartmann, & E. Kris (Eds.), *The psychoanalytic study of the child*. New York: International Universities Press, 1949. Pp. 121-140.
12. LEWIN, K. *Dynamic theory of personality*. New York: McGraw-Hill, 1935.
13. MAIER, N. *Frustration: the study of behavior without a goal*. New York: McGraw-Hill, 1949.
14. MASLOW, A. H. Conflict, frustration and the theory of threat. *J. abnorm. soc. Psychol.*, 1943, 38, 81-86.
15. MILLER, N. E., & DOLLARD, J. *Personality and psychotherapy*. New Haven: Yale Univer. Press, 1941.
16. MILLER, N. E., SEARS, R. R., MOWRER, O. H., DOOB, L. W., & DOLLARD, J. The frustration-aggression hypothesis. *Psychol. Rev.*, 1941, 48, 337-342.
17. MOWRER, O. H., & VIEH, P. An experimental analogue of fear from a sense of helplessness. *J. abnorm. soc. Psychol.*, 1948, 43, 193-200.
18. ROSENZWEIG, S. An outline of frustration theory. In J. McV. Hunt (Ed.), *Personality and the behavior disorders*. New York: Ronald Press, 1944. Pp. 379-388.
19. SEARS, R. R. Nonaggressive reactions to frustration. *Psychol. Rev.*, 1941, 48, 343-345.
20. WINKLEY, R., & JACKSON, KATHERINE. Emotional reactions and behavior of children in the home. *J. Pediat.*, 1951, 38, 476-481.

PARENTAL CORRELATES OF CHILD'S IQ AND HEIGHT: A CROSS-VALIDATION OF THE BERKELEY GROWTH STUDY RESULTS¹

JEROME KAGAN and HOWARD A. MOSS*

Fels Research Institute

In recent years there has been increasing interest in the degree to which environmental pressures and events can influence both IQ level and IQ change during the childhood years (2, 6, 8, 9, 11). Most investigators have reported an average correlation of .50 between the IQ of the school age child and the IQ of the parent (3, 5, 10), and this finding has often been interpreted as indicating the influence of genetic mechanisms on intelligence test performance. However, the size of these correlations leaves much variance to be accounted for, and there are data to suggest that environmental variables may also influence IQ score. Eels and Davis (6) have reported positive correlations ranging from .20 to .43 between socioeconomic level and various IQ test scores. One interpretation of these correlations is that the child in a middle- or upper-class home is more apt to encounter verbal stimulation and encouragement of language skills than a child from a lower class family. It is assumed that this middle class emphasis on language development facilitates IQ score. Kagan and Moss (8) have found that parental attitudes do influence early IQ, for ratings of maternal concern with intellectual development were positively correlated with the IQ of the 3-year-old child when maternal IQ was held constant. However, the controversy over whether IQ is determined by nature or nurture seems less intense now than it was a decade ago. Psychologists are beginning to address themselves to a more plausible question: namely, to what degree and in what ways do genetic and environmental variables influence IQ score.

* Fels Research Institute, Yellow Springs, Ohio.

¹ This research was supported by research grant M-1260 from the National Institute of Mental Health of the National Institutes of Health, U. S. Public Health Service. The authors wish to thank Dr. Nancy Bayley for her reading of the manuscript.

CHILD DEVELOPMENT

In a fairly recent study of parent-child similarities, Bayley (1) reported that as the child grew older there was an increasing degree of similarity between parent and child in height, weight, and intelligence. Bayley used the Berkeley Growth Study population as her source of information and she concluded her paper with the following statement:

Although the sample here is too small to allow for anything but tentative interpretations, it seems very likely that the changing parent-child relationships exhibited here are widespread. They deserve more extensive investigation in other populations and in other characteristics (1, p. 20).

The present report summarizes similar findings for the Fels longitudinal population which is a predominantly middle-class sample and one which seems quite comparable to the Berkeley group. The purpose of this paper is not only to compare the Fels and Berkeley data but also to offer evidence suggesting the influence of maternal concern with intellectual skills on the IQ of the child.

Bayley found that from birth to age 10 there was an increasing and significant correlation between the height of the child and the height of the same sex parent.² The degree of relationship between the child's height and the height of the opposite sex parent was also positive but not statistically significant during these first 10 years of growth. However, the correlations between the child's mental test score³ and the educational level of each parent (Bayley used level of education as an index of parental intelligence) showed a slightly different pattern, for the child's IQ did not show a preferentially high correlation with the education of the same sex parent. Table 1 summarizes Bayley's data for these two variables.

The thing to note in Table 1 is that the correlation between mother's education and child's IQ was slightly higher at ages 3 and 6 for both sexes than the corresponding correlation between father's education and child's IQ. This finding is in contrast to the height data for which the correlations with the same sex parent are larger than the correlations with the opposite sex parent. In addition, for ages 6 and 10 the correlation between mother's education and child's IQ was slightly larger for girls (.61, .51) than it was for boys (.47, .38). These results suggested that maternal concern with intellectual skills played a special role in the development of the child's intellectual potential, and this hypothesis was a major stimulus for the present analysis of the Fels material.

² Bayley's data deal with the correlations between parent and child up to age 21 for height and up to age 19 for mental test score. The Fels data to be reported deal only with the first 10 years of the child's life, and the present summary of Bayley's results refers only to her findings for the first 10 years of life. The reader is referred to Bayley's original paper for additional information.

³ The child's mental test scores were based on the California First Year Scale from 1 to 15 months of age, the California Preschool Scale from ages 1½ through 5, and the Stanford-Binet from ages 6 through 12.

TABLE I

CORRELATIONS BETWEEN (a) HEIGHTS OF PARENTS AND CHILDREN AND
(b) PARENTS' EDUCATION AND CHILDREN'S MENTAL
TEST SCORES (AFTER BAYLEY)

Age	N	B O Y S				G I R L S				
		Height		Education		Height		Education		
		Fa	Mo	Fa	Mo	N	Fa	Mo	Fa	Mo
1 mos. . . .	24	.30	.05	-.29	.00	26	.36	.30	-.14	-.05
3 yrs. . . .	25	.45*	.25	.39*	.49**	24	.38	.67**	.26	.31
6 yrs. . . .	24	.52**	.35	.34	.47*	24	.26	.75**	.56**	.61**
10 yrs. . . .	23	.60**	.37	.44*	.38	24	.24	.72**	.58**	.51**

* Significant at .05 level for two tails.

** Significant at .01 level for two tails.

METHOD

The data to be reported are for a sample of Fels children (59 boys and 40 girls) for whom the following information was available: (a) Stanford-Binet IQ scores from ages 3 through 11; Forms L and M of the test were given alternately, semiannually from the years 2½ to 6 and annually from ages 6 through 11; (b) Otis IQ scores for the majority of the mothers and fathers of the children; (c) level of education for mother and father; (d) heights of parents and heights of the children at birth and at ages 3, 6, and 10. The height data were available for a larger sample of subjects.

Stanford-Binet IQ Scores

Because of the error variance associated with any single IQ score, smoothed IQ scores were obtained for ages 3, 6, and 10 by averaging, for each subject, the three IQ scores around each of these three ages. The rationale for this procedure is described in a recent monograph (11). Thus, the smoothed IQ for age 3 was obtained by averaging the scores for ages 2½, 3, and 3½. The smoothed IQ at age 6 was obtained by averaging the three IQ scores at ages 5½, 6, and 7, and the smoothed IQ at age 10 was obtained by averaging the three IQ scores at ages 9, 10, and 11. Finally, the smoothed IQ scores were normalized using McCall's T-score technique (7).

Parental IQ and Education

The Otis IQ scores were available for most of the mothers and the fathers in this group. The mothers were given Form A and the fathers Form B, and these scores were also normalized using McCall's T-score technique.

Level of parental education was scored by assigning a score of 1 through 6 to each parent according to the following scheme:

CHILD DEVELOPMENT

1. Eighth grade or less.
2. Part high school.
3. High school graduate.
4. Part college.
5. College graduate.
6. Post graduate training.

Table 2 shows the distribution of education levels for the parents of the children used in this study.

TABLE 2
DISTRIBUTION OF EDUCATIONAL LEVEL AMONG THE FELS PARENTS

	B O Y S		G I R L S	
	<i>Father</i>	<i>Mother</i>	<i>Father</i>	<i>Mother</i>
	Per Cent	Per Cent	Per Cent	Per Cent
Eighth grade or less	0.0	0.0	0.0	0.0
Part high school	5.1	2.6	7.1	3.6
High school graduate	33.3	30.8	32.1	28.6
Part college	10.3	35.9	14.3	42.9
College graduate	51.3	30.8	46.4	25.0
Post graduate training	0.0	0.0	0.0	0.0

The average level of education was 4.0 for the fathers and 3.9 for the mothers. These data are comparable to the parental education levels in Bayley's sample, for the average number of years of schooling in her group was 13.7 and 12.8 for the fathers and mothers, respectively. Product-moment correlations were obtained relating (a) parental education and parental IQ with the child's smoothed IQ at ages 3, 6, and 10 and (b) height of each parent with the child's height at birth and at ages 3, 6, and 10.

TABLE 3
CORRELATIONS BETWEEN CHILD'S SMOOTHED IQ SCORE AND
LEVEL OF PARENTAL EDUCATION

<i>Age</i>	B O Y S				G I R L S			
	Parental Education				Parental Education			
	<i>N</i>	<i>Father</i>	<i>N</i>	<i>Mother</i>	<i>N</i>	<i>Father</i>	<i>N</i>	<i>Mother</i>
3 yrs.	59	.17	59	.32*	40	.26	40	.64**
6 yrs.	52	.29*	52	.43**	39	.45**	39	.57**
10 yrs.	46	.24	46	.39**	36	.49**	36	.66**

* Significant at .05 level for two tails.

** Significant at .01 level for two tails.

TABLE 4
CORRELATIONS BETWEEN CHILD'S SMOOTHED IQ SCORE
AND PARENTAL IQ

Age	B O Y S				G I R L S			
	Parental IQ				Parental IQ			
	N	Father	N	Mother	N	Father	N	Mother
3 yrs.	50	.41**	59	.37**	37	.32*	40	.59**
6 yrs.	46	.39**	52	.42**	36	.57**	39	.40**
10 yrs.	41	.30*	46	.41**	33	.51**	36	.48**

* Significant at .05 level for two tails.

** Significant at .01 level for two tails.

RESULTS

Tables 3 and 4 show the product-moment correlations between the child's IQ for ages 3, 6, and 10 and the parental education and IQ scores.⁴

The first thing to note is that the correlation between maternal education and child's IQ is higher than the correlation between paternal education and child's IQ at all ages. For girls at age 3, the difference between the size of the maternal and paternal correlations is statistically significant ($p < .05$, two tails).

Second, maternal education was more highly correlated with the girls' IQ scores than it was with the IQ of boys at all three ages. The differences in the sizes of the maternal correlations with boys' IQ versus girls' IQ at ages 3 and 10 were statistically significant ($p < .05$, $< .10$, respectively, two tails). These data show some similarity to Bayley's results, for she also found that maternal education predicted girls' IQ slightly better than it did boys' IQ for ages 6 and 10.

The correlations between parental IQ and child's IQ score were somewhat different from the correlations which involved parental education. First, the size of the mother-child correlations were not consistently higher than the father-child correlations. Second, unlike the maternal education variable, maternal IQ did not predict girls' IQ significantly better than it did the IQ of boys. It is interesting to note that for the girls maternal education was a slightly better predictor of IQ at all three ages than was maternal IQ.

The Fels data on the height variable are summarized in Table 5.⁵

⁴ The correlation between parental IQ and level of education was .55 for the fathers and .70 for the mothers.

⁵ The authors wish to thank Dr. Stanley M. Garn for the data on parent-child similarities in height.

CHILD DEVELOPMENT

TABLE 5

CORRELATIONS BETWEEN CHILD'S HEIGHT AND HEIGHT OF PARENTS

Age	B O Y S		G I R L S	
	N	Father	N	Mother
Birth	34	.36*	28	.28
3 yrs.	91	.45**	84	.20
6 yrs.	75	.44**	68	.33**
10 yrs.	60	.47**	57	.36**
			49	.34*
			56	.25

* Significant at .05 level for two tails.

** Significant at .01 level for two tails.

Unlike Bayley's data, there is not a general increase in the correlation between the height of the child and that of the same sex parent as the child developed. Rather, paternal height was a better predictor than maternal height of the height of both boys and girls for all ages. The major similarity between the Fels and Berkeley height data is that the correlations between the heights of fathers and sons were higher than the correlations between heights of mothers and sons. We have no ready explanation for these differences. The mean heights of the parents of each sample were quite similar. The average height of the Berkeley parents was 69.7 inches for the fathers and 64.5 inches for the mothers, while the corresponding Fels averages were 69.6 inches and 64.3 inches.

One tentative explanation focuses on the fact that most of the Berkeley group are from an urban population while most of the Fels sample reside in a rural setting. A large majority of the fathers in the Fels sample are or have been engaged in farm labor or other work involving continuous, gross motor activity. It is possible that, if there are selective factors which predispose certain genotypes to enter these vocations, then the Fels sample might be composed of a somewhat different genetic population than the Berkeley group. At present, the most parsimonious hypothesis to account for these differences is in terms of sampling error.

DISCUSSION

The results from both Bayley's work and the Fels material with respect to the child's IQ suggests that (a) maternal education is superior to paternal education in predicting child's IQ at ages 3 and 6 and (b) maternal education predicts girls' IQ better than boys' IQ for ages 6 and 10. In addition, the Fels data suggest that, for girls, maternal education is a somewhat better predictor of IQ than is maternal Otis IQ score.

It is suggested that maternal education is a good predictor of IQ score in both sexes because level of education may be a sensitive index

of the mother's concern with her child's intellectual development. That is, level of maternal education may be positively correlated with (a) the degree to which the mother rewards and encourages intellectual development in the child and (b) the degree to which the child perceives that the mother values intellectual skills and activity. Since the mother's contact with the child during the preschool and early school years tends to be more extensive than that of the father, it is plausible to expect that maternal concern with intellectual development will have a stronger influence on the child's desire to master intellectual skills than similar concern on the part of the father. The fact that maternal education predicts girls' IQ better than boys' IQ might be anticipated since a daughter is more apt than a son to identify with the values of the mother. Previous research by the authors (9) tends to support these hypotheses, for ratings of maternal concern with intellectual development during the child's first three years showed a significant, positive relationship, for girls only, with both achievement fantasy and increase in IQ during the years 6 through 10.

Although these results tentatively suggest that environmental factors may play a role in the determination of intelligence test performance, it is not implied that genetic mechanisms do not influence the IQ. Cyril Burt recently stated,

... with a few rare exceptions, like eye colour or serological difference in the blood, every observable characteristic that geneticists have studied has proved to be the product of the joint action of both heredity and environment. There are, in short, no such things as hereditary characters: there are only hereditary tendencies (4, p. 6).

As Burt states the case, the role of research is to discover the amount of variance that can be attributed to each of these two broad classes of variables. Research is now in progress at this Institute which is directed at a more refined analysis of the various mechanisms which may contribute to intelligence test performance.

SUMMARY

In a previously reported paper Bayley found that the relationship between (a) parental height and child's height and (b) parental education and child's mental test score increased with age. However, the pattern of correlations between parental education and the child's mental score differed from the parent-child height correlations. The height variable showed higher correlations between the child and the same sex parent than between the child and the parent of the opposite sex. The correlations between parental education and child's IQ tended to be higher for maternal education than for paternal education for the preschool years.

The present paper summarized similar data for the Fels research population. For a sample of 59 boys and 40 girls, product-moment correlations were computed between child's smoothed Stanford-Binet IQ at ages 3, 6,

CHILD DEVELOPMENT

and 10 and (a) parental education and (b) parental Otis IQ score. The results revealed that maternal education showed higher correlations than paternal education with the IQ of both boys and girls at all three ages. Second, maternal education was a better predictor of the IQ of girls than of boys for all ages. It was suggested that the results of Bayley and those of the present study indicate that environmental pressures (in the form of maternal acceleration of intellectual growth) may play an important role in the determination of the child's IQ score.

In addition, the Fels data on parent-child similarities in height did not completely agree with those reported by Bayley. The Fels data indicated that paternal height showed higher correlations than maternal height with the height of both boys and girls during the first 10 years of life. There seems to be no ready explanation for this difference between the Berkeley and Fels material.

REFERENCES

1. BAYLEY, NANCY. Some increasing parent-child similarities during the growth of children. *J. educ. Psychol.*, 1954, 45, 1-21.
2. BAYLEY, NANCY, & JONES, H. E. Environmental correlates of mental and motor development: A cumulative study from infancy to six years. *Child Developm.*, 1937, 4, 329-341.
3. BURKS, B. S. The relative influence of nature and nurture upon mental development: A comparative study of foster parent-foster child resemblance and true parent-true child resemblance. *Yearb. Nat. Soc. Study Educ.*, 1928, 27(I), 219-316.
4. BURT, C. The inheritance of mental ability. *Amer. Psychologist*, 1958, 13, 1-15.
5. CONRAD, H. S., & JONES, H. E. A second study of familial resemblances in intelligence: Environmental and genetic implications of parent-child and sibling correlations in the total sample. *Yearb. Nat. Soc. Study Educ.*, 1940, 39(II), 97-141.
6. EELS, K., DAVIS, A., et al. *Intelligence and cultural differences: a study of cultural learning and problem solving*. Chicago: Univer. of Chicago Press, 1951.
7. JOHNSON, P. O. *Statistical methods in research*. New York: Prentice Hall, 1949.
8. KAGAN, J., & MOSS, H. A. Maternal influences on early IQ scores. *Amer. Psychologist*, 1958, 13, 320. (Abstract)
9. KAGAN, J., & MOSS, H. A. The stability and validity of achievement fantasy. *J. abnorm. soc. Psychol.*, 1959, 58, 357-364.
10. LEAHY, A. M. Nature-nurture and intelligence. *Genet. Psychol. Monogr.*, 1935, 17, 235-308.
11. SONTAG, L. W., BAKER, C. T., & NELSON, VIRGINIA L. Mental growth and personality development: a longitudinal study. *Monogr. Soc. Res. Child Developm.*, 1958, 23, No. 2. (Serial No. 68)

RELATIONSHIPS AND AGE DIFFERENCES IN GROWTH MEASURES AND MOTOR SKILLS¹

LOUIS A. GOVATOS*

Western Michigan University

Considerable emphasis has been placed on the study of motor performances of children (1, 2, 5, 6, 9, 10). Most studies² have been concerned primarily with achievement, ability, and developmental aspects of motor performances at the preschool, elementary, and secondary levels. In the field of physical education much research has been concerned with establishing normative data, increasing efficiency in motor skills by practice, and devising classification systems based on factors of height, weight, and age of participants. Some treatment has been given to establishing relationships between various growth measures and achievement in certain physical skills. A cross-sectional study by Seils (8) showed relationships existing between specific physical skills, such as the 40-yard dash, standing broad jump, and throwing for distance, and various growth measures, such as height, weight, and carpal development of children in the primary grades. However, very little emphasis has been given to studying the relationships existing between performances in specific gross motor skills and a comprehensive assessment of the child's developmental status.

STATEMENT OF THE PROBLEM

The specific purposes of this study were:

1. To determine what relationships exist among selected measures of growth of the elementary school child in size, bone structure, strength, intelligence, and performance in certain motor skills involving running, jumping, throwing, and kicking.
2. To determine what differences exist between boys and girls in the various growth measures and motor skills.

* School of Education, Western Michigan University, Kalamazoo.

¹ This study was part of a longitudinal, multidiscipline investigation completed as a doctoral dissertation in 1952 at the University of Michigan with Willard C. Olson as advisor.

² For a more comprehensive review see "Relationship between Physical Skills and Growth in Elementary School Children," unpublished doctoral dissertation, University of Michigan, Ann Arbor, 1952.

CHILD DEVELOPMENT

DESIGN OF THE STUDY

Out of a total enrollment of 137 children attending the Elementary School of the University of Michigan, only 101 were selected for this study. Their ages ranged from 6 to 11 years. For these subjects complete information was available on the following growth variables: height, weight, grip, dental, carpal, reading and mental ages. The remaining children were excluded because of incomplete growth records, illness, and other physical handicaps. The average IQ was approximately 125 for this group of children. Most came from families of moderate to above average socioeconomic status. It is apparent that this sample constitutes an atypical group and therefore the findings cannot be generalized to the total school population.

The growth measures consisted of height expressed in inches, weight in pounds, strength of grip in kilograms, and the number of permanently erupted teeth. In order to express these attributes of growth in similar units of measurement, these measures were converted into age values (7). In addition to the above measures of growth, age values were obtained for carpal development by assessing X-rays of the hand and wrist with typical age samples found in Flory's standards (3). Similarly, intelligence and reading tests yielded mental and reading ages, respectively. An organismic age was established for each subject by taking an average of the growth ages.

The motor skills included the 25-yard dash, soccer kick for distance, softball throw for distance (over- and underhand), softball throw for accuracy (over- and underhand), standing broad jump, and the jump and reach. The raw motor skill scores were converted into age skill scores by establishing a regression line by the least squares method (4).

For the purposes of correlation the boys, within various age groupings, were combined into one group. Likewise, the girls were combined in the same manner with separate treatment given to each sex. Pearson's product-moment correlations and partial correlations were used to describe the relationships between the growth measures and the motor skills. For differences in the mean growth ages and mean performances in motor skills for the boys and girls, t values were obtained.

RESULTS OF THE STUDY

Tables 1 and 2 present t values for differences in the mean motor performances and mean growth ages, respectively, for elementary school children. The remaining tables contain correlations between the growth measures and motor skills for both boys and girls.

In Table 1 significant differences are recorded in the mean motor performances of boys and girls within the same chronological age groupings. As indicated by their mean performance on the standing broad jump, the 6-year-old girls were superior to the boys of comparable ages. Where signifi-

LOUIS A. GOVATOS

TABLE 1
DIFFERENCES IN MOTOR SKILLS FOR BOYS AND GIRLS
CLASSIFIED IN 12-MONTH AGE INTERVALS

AGE INTERVAL:	66-77	78-89	90-101	102-113	114-125	126-137
SEX:	Boys Girls	Boys Girls	Boys Girls	Boys Girls	Boys Girls	Boys Girls
N:	8 3	12 10	6 7	12 13	11 6	7 6
	t	t	t	t	t	t
Jump and Reach ..	+0.40	+0.41	-0.37	-1.16	+0.54	+0.67
Standing Broad						
Jump	+2.42*	-1.16	-0.20	-1.22	+0.48	-0.92
Soccer Kick						
for Distance	-0.52	-3.30**	-1.33	-1.93	-2.31*	-1.52
Twenty-five-Yard						
Dash	+1.20	+0.06	-1.30	-0.71	+0.65	-0.00
Ball Throw for Distance (Overhand)	-1.80	-4.61**	-2.54*	-0.40	-2.35*	-6.72**
Ball Throw for Distance (Underhand)	-0.87	-4.93**	-2.37*	-2.69*	-0.85	-2.41*
Ball Throw for Accuracy (Overhand)	-1.36	-3.61**	-2.23	-2.11*	-1.08	-3.28**
Ball Throw for Accuracy (Underhand)	-1.06	-1.73	-3.17**	-2.66*	-0.14	-0.45**

NOTE.—A minus sign indicates a better mean performance for the boys. A plus sign indicates a better mean performance for the girls.

* Significant at 5 per cent level.

** Significant at 1 per cent level.

TABLE 2
DIFFERENCES IN GROWTH AGES OF BOYS AND GIRLS
CLASSIFIED IN 12-MONTH AGE INTERVALS

AGE INTERVAL:	66-77	78-89	90-101	102-113	114-125	126-137
SEX:	Boys Girls	Boys Girls	Boys Girls	Boys Girls	Boys Girls	Boys Girls
N:	8 3	12 10	6 7	12 13	11 6	7 6
	t	t	t	t	t	t
Height Age	-0.62	-2.32*	+1.04	+0.15	-0.61	-0.20
Weight Age	-0.58	-2.06	-0.10	+0.08	-1.59	-0.63
Grip Age	+0.16	-0.80	+0.13	-0.74	-0.77	-1.21
Carpal Age	-0.74	-2.17*	-0.35	+0.84	-1.30	-0.45
Dental Age	+2.34*	-1.16	+0.20	-0.47	-1.19	+0.29
Reading Age	-0.36	-1.43	+2.29*	+3.59**	-0.03	+1.95
Mental Age	-0.07	-2.00	+2.11	+2.40*	+0.29	+0.99
Organismic Age ..	-0.41	-3.17**	+0.69	+1.59	-0.89	+0.30

NOTE.—A minus sign indicates a better mean performance for the boys. A plus sign indicates a better mean performance for the girls.

* Significant at 5 per cent level.

** Significant at 1 per cent level.

CHILD DEVELOPMENT

cant differences were shown at succeeding age levels, the boys excelled the girls in mean performances on the soccer kick for distance, ball throw for distance (over- and underhand), and ball throw for accuracy (over- and underhand).

Few significant differences in the mean growth ages of boys and girls within various age groupings were found. In Table 2 the 48 values expressed between the growth measures and the various age groupings revealed only seven significant differences. At the 6-year level the girls were superior to the boys in mean dental development. In the 7-year-old group the boys exceeded the girls in mean height, carpal, and organismic ages. The girls excelled the boys in mean reading age at both the 8- and 9-year level. There were no significant differences found between boys and girls in any of the growth measures at the 10- and 11-year intervals.

TABLE 3
CORRELATIONS BETWEEN MOTOR SKILLS AND VARIOUS ASPECTS OF
GROWTH IN ELEMENTARY SCHOOL BOYS AND GIRLS

Motor Skills		ASPECTS OF GROWTH						
		Height	Weight	Grip	Carpal	Dental	Reading	Mental
Jump and Reach	(B) ..	.55	.42	.63	.50	.63	.45	.66
	(G) ..	.67	.53	.66	.70	.69	.54	.51
Standing Broad Jump	(B) ..	.60	.44	.71	.55	.70	.54	.71
	(G) ..	.62	.51	.63	.69	.72	.65	.65
Soccer Kick for Distance	(B) ..	.70	.68	.75	.70	.80	.61	.66
	(G) ..	.73	.64	.71	.78	.77	.63	.66
Twenty-five-Yard Dash	(B) ..	.53	.41	.59	.46	.57	.38	.61
	(G) ..	.59	.41	.52	.62	.63	.50	.51
Ball Throw for Distance (Overhand)	(B) ..	.70	.64	.84	.66	.71	.55	.65
	(G) ..	.68	.60	.71	.77	.77	.59	.60
Ball Throw for Distance (Underhand)	(B) ..	.67	.63	.77	.63	.70	.55	.64
	(G) ..	.79	.68	.75	.77	.80	.57	.64
Ball Throw for Accuracy (Overhand)	(B) ..	.60	.54	.63	.47	.56	.40	.53
	(G) ..	.48	.35	.46	.62	.51	.45	.46
Ball Throw for Accuracy (Underhand)	(B) ..	.57	.54	.58	.60	.61	.42	.55
	(G) ..	.58	.52	.67	.63	.65	.53	.56

NOTE.—All the above correlations were significant at the 1 per cent level excepting ball throw for accuracy (overhand) with weight, which, for the girls, was significant at the 2 per cent level.

In Table 3 it may be noted that all of the correlations between the growth measures and performances in motor skills were significant at the 1 per cent level, with the exception of the overhand ball throw for accuracy with weight for the girls. For the boys the soccer kick for distance correlated highest with dental age (.80), and lowest with reading age (.61). The overhand throw for distance correlated highest with grip age (.84), and lowest with reading (.55). The lowest relationships were found between

LOUIS A. GOVATOS

reading age and the 25-yard dash (.38) and the overhand ball throw for accuracy (.40). For the girls the highest relationships existed between dental age and the underhand ball throw for distance (.80), and carpal age with soccer kick for distance (.78). The lowest relationships were found to be weight age with the 25-yard dash (.41) and overhand ball throw for accuracy (.35).

TABLE 4

PARTIAL CORRELATIONS (HOLDING CHRONOLOGICAL AGE CONSTANT)
BETWEEN VARIOUS ASPECTS OF GROWTH AND MOTOR SKILLS
IN ELEMENTARY SCHOOL BOYS AND GIRLS

Motor Skills		ASPECTS OF GROWTH						
		Height	Weight	Grip	Carpal	Dental	Reading	Mental
Jump and	(B) ..	.03	-.02	.18	-.12	.01	-.01	.19
Reach	(G) ..	.17	.03	.32*	.25	.30*	-.07	-.24
Standing	(B) ..	.01	-.09	.23	-.19	.02	.06	.14
Broad Jump	(G) ..	-.03	-.07	.23	.14	.33*	.14	.05
Soccer Kick for	(B) ..	.28*	.41**	.38**	.25	.42**	.22	.05
Distance	(G) ..	.34*	.26	.40**	.46**	.48**	.14	.13
Twenty-five-Yard	(B) ..	.00	-.03	.10	-.20	-.13	-.12	.07
Dash	(G) ..	.14	-.09	.14	.19	.29	-.00	-.05
Ball Throw for Dis-	(B) ..	.27*	.33*	.59**	.13	.10	.09	.02
tance (Overhand)	(G) ..	.00	.06	.35*	.26	.40**	-.15	-.23
Ball Throw for Dis-	(B) ..	.29*	.33*	.47**	.15	.20	.16	.09
tance (Underhand)	(G) ..	.35*	.25	.43**	.28	.47**	-.21	-.11
Ball Throw for Accu-	(B) ..	.31*	.28*	.37**	.04	.14	.05	.11
racy (Overhand)	(G) ..	-.11	-.18	.05	.22	.05	-.11	-.14
Ball Throw for Accu-	(B) ..	.21	.26	.20	.24	.20	.04	.09
racy (Underhand)	(G) ..	.01	.04	.37*	.13	.26	-.05	-.04

* Significant at 5 per cent level.

** Significant at 1 per cent level.

In order to determine whether chronological age had influenced the high correlations obtained in Table 3, partial correlations were calculated holding chronological age constant. Analyses of this nature have the effect of depressing the magnitude of the correlation values. Many of the relationships between growth measures and the motor skills were found still to be significant (see Table 4). For the boys the soccer kick for distance correlated significantly with height (.28), weight (.41), grip (.38), and dental age (.42). Significant correlations were also obtained for grip age with the overhand ball throw for distance (.59), underhand ball throw for distance (.47), and overhand ball throw for accuracy (.37). For the girls, the soccer kick for distance correlated significantly with grip (.40), carpal (.46), and dental age (.48). There was a significant relationship between the underhand ball throw for distance and height (.35), grip (.43), and dental age (.47). For neither boys nor girls did reading and mental ages correlate significantly with any of the motor skills.

CHILD DEVELOPMENT

TABLE 5

TOTAL AND PARTIAL CORRELATIONS (HOLDING CHRONOLOGICAL AGE CONSTANT) BETWEEN ORGANISMIC AGE AND MOTOR SKILLS IN ELEMENTARY SCHOOL BOYS AND GIRLS

Motor Skills (in Age Units)	ORGANISMIC AGE			
	Total Correlations		Partial Correlations	
	Boys N=56	Girls N=45	Boys N=56	Girls N=45
Jump and Reach64**	.67**	.06	.10
Standing Broad Jump70**	.71**	.04**	.12
Soccer Kick for Distance63**	.77**	-.15	.42**
Twenty-five-Yard Dash59**	.59**	-.05	.07
Ball Throw for Distance (Overhand)79**	.73**	.37**	.07
Ball Throw for Distance (Underhand)76**	.69**	.40**	-.10
Ball Throw for Accuracy (Overhand)62**	.52**	.31*	-.10
Ball Throw for Accuracy (Underhand) ..	.64**	.70**	.28*	.27

* Significant at 5 per cent level.

** Significant at 1 per cent level.

Although the product-moment correlations in Table 5 were significant at the 1 per cent level, when chronological age was held constant some of these relationships between organismic age and the motor skills became insignificant. For example, only the soccer kick for distance correlated significantly with organismic age (.42) for the girls. For the boys, organismic age correlated significantly with the overhand ball throw for distance (.37), underhand ball throw for distance (.40), overhand ball throw for accuracy (.31), and the underhand ball throw for accuracy (.28).

SUMMARY AND CONCLUSIONS

Increases in the mean motor performances for both boys and girls were observable with each advancing age level in this study. Increases in mean motor performances were more evident at various age levels for some of the motor skills measured, for example, for the jump and reach for boys between 8 and 9 years of age, and for girls between 9 and 10. There were marked increases on the overhand ball throw for distance for boys from 6 to 7 and between the 10- and 11-year age levels, and likewise for the girls from the 9- to the 10-year interval. On the ball throw for distance (underhand), the boys showed the greatest increase in achievement between the 6- and 7-year age levels, and the girls from the 9- to the 10-year interval. There was also evidence of sex differences in the mean motor performances and also in the mean growth measures by age intervals. The 6-, 7-, 10-, and 11-year-old girls were superior to the boys of comparable ages on the jump

LOUIS A. GOVATOS

and reach. The boys surpassed the girls at the 7-, 8-, 9-, and 11-year levels on the standing broad jump. At each age interval the boys were superior to the girls on the soccer kick, over- and underhand ball throw for distance, and the over- and underhand ball throw for accuracy. However, the girls displayed superiority in the 25-yard dash at the 6-, 7-, and 10-year age levels.

When the individual aspects of growth were correlated with each motor skill, the resultant relationships were found to be high and positive for both boys and girls in this study. However, when chronological age was held constant, the resulting relationships between each growth measure and the individual motor skill were found to be low and in some cases negative.

Although the evidence presented thus far in this study is based upon a limited sample, the following conclusions seem warranted:

1. When the individual aspects of growth and organismic age were correlated with each motor skill, the resulting relationships were found to be positive and significant for both boys and girls. This would seem to indicate that gross physical development is accompanied by the ability to manipulate the parts of the body that have grown.

2. In Table 3 the jump and reach and the 25-yard dash tend consistently to exhibit higher correlations with some measures of growth for the girls than for the boys. However, this consistency does not hold for other motor skills. As indicated in Table 1, this apparent advantage for the girls may be related to their earlier physiological development and also to their play activities involving games similar to hop-scotch, skipping rope, and playing tag. Boys tend to excel in activities involving strength and endurance, for example, in games requiring both kicking and running.

3. When chronological age was held constant (Tables 4 and 5), the resulting relationships between each growth measure and between organismic age and each motor skill were found to be low and in some cases negative. However, both dental and grip age revealed significant relationships with jump and reach, soccer kick for distance, and over- and underhand ball throw for distance. Apparently these growth measures appear to be better indices than height and weight for predicting success in some of the motor skills studied. None of the growth measures was related significantly to the 25-yard dash. Organismic age was found to be significantly related at the 1 per cent level of confidence only to soccer kick for the girls and to the over- and underhand ball throw for distance for the boys. The apparent lack of relationship between organismic age and most of the motor skills suggests the inadvisability of using a single battery for establishing standards of expectancies for groups of children at a particular age level.

4. Both reading and mental growth measures approach zero correlations with the motor skills indicated in Table 4. Apparently mental growth measures are not as important as physical attributes in predicting performances in gross motor skills. Rather their relationships seem to be greater with motor skills involving detailed directions, finer motor coordinations, and frequent changes in movement.

CHILD DEVELOPMENT

5. In many cases boys and girls of similar growth status seem to be equally effective in activities involving running and jumping. However, boys appear to excel girls in throwing and kicking. From a curriculum point of view there are strong indications that boys and girls may engage in joint activities that involve a combination of many motor skills.

REFERENCES

1. CARPENTER, A. Tests of motor educability for the first three grades. *Child Develpm.*, 1940, 11, 293-299.
2. CUNNINGHAM, B. V. An experiment in measuring gross motor development of infants and young children. *J. educ. Psychol.*, 1927, 18, 458-464.
3. FLORY, C. D. Osseous development in the hand as an index of skeletal development. *Monogr. Soc. Res. Child Develpm.*, 1936, 1, No. 3.
4. GUILFORD, J. P. *Fundamental statistics in psychology and education*. New York: McGraw-Hill, 1950.
5. JENKINS, L. M. *A comparative study of motor achievements of children of five, six and seven years of age*. New York: Teachers Coll., Columbia Univer., Bureau of Publications, 1930.
6. JONES, H. E. Relationships in physical and mental development. *Rev. educ. Res.*, 1936, 6, 102-123.
7. OLSON, W. C., & HUGHES, B. O. Manual for the description of growth in age units. Ann Arbor, Mich., 1950.
8. SEILS, L. G. The relationship between measures of physical growth and gross motor performance of primary school children. Unpublished doctoral dissertation, Boston Univer., 1948.
9. WARNER, A. P. The motor ability of third, fourth, and fifth grade boys in the elementary school. Unpublished doctoral dissertation, Univer. of Michigan, 1952.
10. WELLMAN, B. L. Physical growth and motor development and their relation to mental development in children. In C. Murchison (Ed.), *A handbook of child psychology*. Worcester, Mass.: Clark Univer. Press, 1931. Pp. 242-277.

INTERVIEWS OF PARENTS OF HIGH ANXIOUS AND LOW ANXIOUS CHILDREN¹

KENNETH S. DAVIDSON*

Yale University

An important part of an ongoing project concerned with anxiety in children was a series of interviews with the parents of a selected group of high anxious (HA) and low anxious (LA) children. Descriptions (5) have been presented of the test anxiety (TA) and general anxiety (GA) scales² which were used to determine placement of children in the LA and HA groups in the present study. In other reports dealing with these same subjects results were presented of their Rorschach (4), Draw-a-Person (2), and learning task³ (6) performances, of observations of their classroom behavior (3), and of their parents' ratings (1) of them on a personality check list.

The previous studies provided evidence bearing on the validity of the two anxiety scales. The interviews of the parents were done to assess the validity of the scales further. However, an additional purpose was to obtain information about the children and their environments that would aid in understanding the development of anxiety in children and in deciding on

* Department of Psychology, Yale University, New Haven, Connecticut.

¹ The project of which this study is a part is being supported by a grant from the U. S. Public Health Service, (M-712).

² Copies of these two scales are available for research purposes only.

³ In the study involving the experimental learning task, 24 of the 32 matched pairs of children were used for reasons cited in that publication.

CHILD DEVELOPMENT

the most fruitful approaches in future investigations. Therefore, for exploratory purposes, many items were formulated and integrated into two standardized interviews without specific predictions about the nature of results or differences between the HA and LA subjects. In a general way, it was predicted that the interview data would describe the LA children more favorably and with fewer and less severe stresses, emotional disturbances, conflicts, and anxieties than the HA children.

Two standardized interviews were developed to provide relevant information with respect to the two anxiety scales. The first interview was focused primarily on the child's school experiences and was more related to the TA scale. The second interview inquired about the subject's personality and emotional development so that information pertaining more to the GA scale was obtained.

METHOD

Subjects

The 64 subjects of this study were chosen from 747 pupils in grades 2 through 5 who had been administered the TA and GA scales the year before. All had taken the Otis Alpha Intelligence Scale and children were excluded who were regarded as serious arithmetic, reading, or behavior problems by their teachers. Additionally, only children whose parents were living and not divorced or separated were selected. Subjects were placed in the HA group if their TA and GA scores fell in the upper quartiles of both distributions; those whose scores were in the lower quartiles were regarded as LA. By matching LA with HA on grade, sex, and Otis Alpha T scores, 32 matched pairs of children were obtained. For each grade-sex group, there were four matched pairs, and in a 2 (anxiety) by 2 (sex) by 4 (grade) arrangement there were four children in each of 16 cells in the design.

Procedure

The parents of each of the 64 subjects were interviewed twice in their homes at their convenience by the author who did not know whether they were parents of a LA or HA subject. In most instances, only the mother was interviewed, although 21 of the fathers were present for all or part of either or both interviews. After all parents had been seen for the first interview, the second interviews were begun. An interval of about two months lapsed before the second interview in each case. After the first interview the mothers completed an attitude scale on child rearing, and in the second interview they rated their children on a personality check list (1). Fathers who were not present completed these rating scales which were left for them. The cooperation of the parents generally can be characterized by the fact that no parent refused to answer any question in either interview.

KENNETH S. DAVIDSON

RESULTS

The analysis of the interview data fell into four categories: biographical data of the children, biographical data of the parents, the childrens' school experiences, and their personality development and emotional adjustment. However, in considering the results in these categories, it was felt that a fifth set of analyses—on the parents' interview behavior—might have important implications. The reason for this was the previous finding (1) that, on the personality check list, fathers of LA rated them significantly more favorably than fathers of HA rated their children while the ratings by the mothers did not differentiate in this respect between the LA and HA subjects. Since the interview data came almost exclusively from mothers, significant differences might be masked, presumably as a result of the mothers of the HA children giving more favorable descriptions of them than the fathers of those children might report. For these reasons the interview behavior of the parents was analyzed and the results will be presented and discussed. Following that, the findings will be presented for the four categories of analyses indicated above. In all the categories the analyses utilized a one-tail test of significance whenever the direction of the results were predicted.

Parental Interview Behavior

The problem that parents' evaluations might mask differences between the LA and HA subjects came to our attention only after the data were collected. As a result, there was no carefully planned attention to, or measurement of, defensiveness during the interviewing. However, after each interview, the interviewer had written his spontaneous reactions—approximately one and a half typed pages—to the mother and the interview. It was decided to evaluate these reports for possible differences between mothers of LA and HA in the tendency to distort or be defensive in answering questions. Such distortions by mothers can be of at least two kinds: conscious withholding or falsification and unconscious defense against recognition both of the mother's and the child's "unacceptable" covert and overt behavior. Of these two kinds of distortions, it was likely that it was the conscious type which may have been noted by the interviewer. Therefore, six judges independently read the two write-ups for each parent with the instructions to rate a parent as "defensive" if the interviewer said something in his report to indicate the parent was consciously censoring the information she was revealing. The judgment was whether *the interviewer had concluded* the parent was defensive, not whether the judge so concluded. Parents not noted as defensive automatically were rated as nondefensive. Judges were unaware of the anxiety group to which a subject belonged. For three of the judges, there was a significant tendency for more mothers of HA boys than mothers of LA boys to be rated as defensive ($p < .10$). For two of the other judges there was a nonsignificant tendency in the

CHILD DEVELOPMENT

same direction while the relationship for the sixth judge was of the chance variety. The ratings of "defensive" were added across judges for each subject and the mothers of the matched pairs of LA and HA boys were compared on this "defensive" score by a dependent t test. The difference between the LA and HA "defensive" scores yielded a t value of 1.745 which indicates a tendency, significant almost at the 5 per cent level, for the HA boys' mothers to get higher "defensive" scores than mothers of the LA boys. The ratings on defensiveness, though in the same direction, did not differentiate significantly between the mothers of the LA and HA girls. Thus, even though the interviewer's write-ups were not as focused on defensiveness as would have been the case if the distortion of the HA mothers had been anticipated, these findings present evidence that the mothers of HA boys were more defensive than the LA boys' mothers.

The implication of this finding is twofold. First, it definitely points to the possibility that the mothers of the HA boys consciously withheld information or distorted the picture in describing their children. For example, the mothers of the HA boys said their children had no more of the anxieties on a check list than the mothers of LA said their boys had, despite the fact that the HA children have admitted to more fears and anxieties. Clearly the suggestion is that the findings for the boys in the categories to be discussed later can be regarded as slanted more favorably toward the HA subjects than would have been the case if their mothers had not been more defensive. Secondly, this finding can be interpreted to mean that HA boys' mothers had more "unacceptable" behaviors to cover up in themselves and in their children, thus lending further evidence to the validity of the anxiety scales. Seeking corroboration of this finding, several objective analyses of the mothers' responses to questions of the interviews were done. All utilized chi square analyses and can be summarized as follows:

1. In a free description of their child's first day of school in the first interview, parents of LA boys used more words (actual count) than parents of HA boys ($p < .01$).

2. To begin the second interview, the examiner said, "Last time we talked about how S felt about school. This time we would like to get a clear picture of what S is like as a person. How would you describe S so that a person like myself who doesn't know S would get an idea of the kind of child S is?" In giving their descriptions, more parents of HA children asked the interviewer a question than parents of the LA subjects ($p < .05$). This finding was more significant for girls ($p < .10$) than boys ($p < .20$).

3. In that free description of their child, parents of HA children said less (word count) before asking the interviewer a question than parents of the LA children ($p < .025$), and they said less in total than LA parents on the same question ($p < .05$).

Thus, it would seem that the LA parents were more spontaneous, told more, and were less dependent on the interviewer than the HA parents in

KENNETH S. DAVIDSON

discussing their children's experiences and personalities. Twice as many parents of HA girls as parents of LA girls failed to mention any affect in telling about their child's first day in school ($p < .10$). This offers slight evidence that the LA girls' mothers tended to deal more freely with emotional factors than the mothers of HA girls. More important, these findings indicate further that HA mothers were more defensive in the interviews than LA mothers. As a result the findings below can be regarded as presenting the HA subjects in a better light than if the HA mother had been less defensive. With these considerations in mind, the remaining analyses will be presented and discussed.

Biographical Data of the Children

The results of the analyses of this data are summarized as follows:

1. HA children had more brothers than sisters, while the LA subjects had more sisters than brothers ($p < .02$). This relationship was more significant for boys ($p < .05$) than for girls.⁴
2. Related to that, more HA than LA subjects did not have any sister ($p < .02$), and this finding was stronger for boys ($p < .10$) than for girls.⁴
3. HA boys had more illnesses than the LA boys ($p < .01$).

Additionally, the LA children tended to be followed by sisters or to be the youngest child, while HA subjects were followed by brothers or are themselves "only" children ($p < .10$). The four only children—three boys and one girl—were all in the HA group. These findings suggest that having sisters was less anxiety inducing, especially for boys, than having brothers.

Biographical Data of the Parents

Several items in the first interview concerned the background of the parents. Of these analyses, the results of those attaining significance are listed here:

1. One question asked whether the parents had ever been away for any reason from S for more than a day. The fathers of LA subjects were away from their children overnight more frequently than the fathers of the HA subjects ($p < .05$). This finding was stronger for boys ($p < .10$) than for girls.⁴
2. The mean level of education of the LA fathers was significantly higher than that of the fathers of the HA children ($p < .01$). This difference, too, was significant for the boys ($p < .01$) but not the girls. The same result is expressed in the finding that more LA fathers graduated from high school than fathers of the HA subjects ($p < .025$), with greater significance again for boys ($p < .05$) than for girls.⁴
3. Then, more mothers of the HA boys worked since the subjects were born than did the mothers of the LA boys ($p < .05$).

⁴ A two-tail test of significance was utilized in obtaining the p values given.

CHILD DEVELOPMENT

Children's School Experiences and Adjustment

This area was of prime importance in view of the focus of this research on "test anxiety." Analyses of interview data yielded the following significant results:

1. Parents said HA children worry more about missing school for illness than LA children do ($p < .02$). This finding was more significant for girls ($p < .025$) than for boys.

2. Another question was "The first day of school is certainly an important one for a child. But it is also an important one for parents. What were your thoughts and feelings about *S* starting school?" Parents of HA children felt more conflict and uncertainty when their children first entered school than the parents of LA subjects ($p < .025$). This finding was also more significant for the boys ($p < .025$) than the girls.

3. HA boys tended to start Sunday School at an older age than LA boys ($p = .055$).⁴

4. The last question of the first interview was "Do you think *S* has learned: (a) *more than* you expected of him, (b) *as much as* you expected of him, (c) *less than* you expected of him?" Parents of LA boys said they learned *as much as* expected while HA boys were described by their parents as learning either *more* or *less* than expected ($p < .05$). In this respect LA girls were said to learn *more* than their parents expected while the parents of HA girls said they learned either *less* than or *as much as* expected of them ($p < .10$).⁴

5. The LA felt more positive than the HA children about starting school this year ($p < .10$) and this was truer for the boys ($p < .05$) than for the girls.

Some trends which approached significance in this category involved communication about school. The LA children found more to tell about school when at home ($p < .10$), and LA boys said more about school before they first entered than HA boys ($p < .10$). The latter subjects had more trouble learning to read than the LA boys ($p < .10$).

In respect to the subjects' school experiences and adjustment, the subjects' actual grades were regarded as important, so they are reported here incidentally, although the data were obtained from school records rather than the interview. The grades were given one full year after the anxiety questionnaires were administered, and matched subjects were compared over-all by dependent t tests and in individual marks by sign tests. The following significant findings, using one-tail tests, were found between the final school marks of the two groups of subjects:

1. LA got higher marks than HA children ($t = 1.900, p < .03$).
2. LA got higher marks than HA children as follows: social studies ($p = .095$); work habits ($p = .039$).
3. LA boys got higher marks than HA boys ($t = 2.331, p < .025$).

KENNETH S. DAVIDSON

4. LA boys got higher marks than HA boys as follows: social behavior ($p = .073$); language arts ($p = .011$); social studies ($p = .090$); science ($p = .090$); music ($p = .035$); and work habits ($p = .055$).

5. LA girls got higher marks than HA girls ($t = .571$, $p < .30$).

Clearly, the LA boys functioned more effectively than their HA mates who were matched on IQ with them. In nine of the 16 matched pairs of LA and HA boys, the LA had better marks, while the HA boys had higher marks in only three. On the other hand, for the girls the results were chance. In half of the pairs the HA had higher grades while the LA girl received higher marks in the other half of the pairs. An interesting point is that whichever girl of a pair did better, tended to do so significantly.

Children's Emotional Adjustment and Development

The major analysis in this category involved the interviewer's write-ups of the first interview. In this, the interviewer rated each sentence of a write-up as to the presence or absence of positive and negative qualities and factors in the environment that pertain to or affect the subjects' emotional adjustment. A positive rating for a sentence would indicate adequate or successful adjustment by the child or an environmental factor promoting such an adjustment. A negative rating of a sentence would indicate difficulties in the child's efforts to cope with his problems. In the event positive and negative aspects were both present and about equal in a sentence, a rating of "both" was given. Sentences which had no bearing on the child's adjustment were rated "neither." To provide a clearer understanding of these ratings, the following are examples of sentences and the judge's evaluations of them.

- + The inner peace and outward quiet is easy to see—they are not demanding or pressing but take what comes and give their children lots of love and affection.
- + Both Mr. and Mrs. ——— seem basically in agreement in basic things, but differences can be tolerated and I feel that whatever problems they have between them—they have few, I would guess—would not be saddled on the children.
- + S is seen as an able, independent child who goes out to meet the world and the experiences it offers.
- S is sensitive and cries easily if reprimanded, punished, frustrated or has difficulties with playmates S can't solve.
- S is a child with phobias and anxieties, some masked, others blatant.
- Mrs. ——— is a nervous woman, expecting danger and trouble everywhere and wants to solve her child's problems before they come up.
- +— But still one gets a feeling of immaturity in many things as well as maturity in many others.
- +— S's feelings can be hurt easily but S is much more resilient than before.

CHILD DEVELOPMENT

- + - She describes *S* as generally meeting situations effectively and eagerly, though she says *S* is slow in school and that *S* worries quite a bit about *S*'s difficulty in keeping up with the class.

As a reliability check, the interviewer's ratings of the sentences were compared with those by an independent judge on the write-ups of the first interviews for eight pairs of LA and HA children. There were four pairs each of boys and girls chosen at random. Each judge independently rated each sentence of 16 reports. When the ratings were completed and compared, it was found that total agreement was present on 263, or 78 per cent, of the 338 ratings each judge had made. Of the 75 disagreements, only four involved completely opposite ratings, plus by one and minus by the other rater. For each of these 16 cases, the number of minus ratings was subtracted from that of the positive ratings and the direction, plus or minus, indicated by sign. In the reliability check this procedure gave two "favorable-unfavorable" scores (one from each judge's ratings) for each subject. The interjudge correlation of these scores was $+ .94$, indicating a high degree of agreement as to over-all favorableness of their ratings for these subjects. Finally, when the dependent t test was applied to the differences in "favorable-unfavorable" scores of the eight matched pairs, the results were significant beyond the 1 per cent level for both judges. That is, both judges rated the LA reports significantly more favorably than those written for the HA subjects.

Then, for each of the 64 subjects the number of plus and minus ratings by the interviewer were added algebraically to obtain a "favorable-unfavorable" score for each child. The differences in these scores for the 32 matched pairs were then subjected to the dependent t test. For all 64 subjects the LA subject write-ups were rated more favorably than those of the HA children ($p < .01$), and the results were significant for both boys ($p < .01$) and girls ($p < .01$). Of all the results obtained from comparisons of LA with HA, this is certainly one of the most striking. Additionally, HA children tended to have more negative than positive statements written whereas the LA tended to have more positive than negative statements written in the reports of the interviews ($p < .025$). That finding was more significant for the boys ($p < .05$) than for the girls ($p < .15$). The interviewer not only did not know at any time which children were HA and which were LA but all through the interviewing had the impression that the write-ups would not predict which were the HA and LA subjects. While these findings are not based on ratings of anxiety, they clearly indicate that the interviewer and an independent judge evaluated the LA children as better adjusted and having more favorable environments in general mental health terms than the HA subjects.

Finally, it was found that LA children tended to have more modes of expressing anger than their HA mates ($p < .025$). For both the girls and the boys separately the findings tended toward significance ($p < .10$). Parents indicated whether or not *S*, when angry, (a) gets sulky and silent,

KENNETH S. DAVIDSON

(b) hits people, (c) throws things, (d) talks angrily, or used any other methods of expressing anger. The suggestion here is that the LA were freer to express aggression, at least in choosing how to vent their feelings and impulses. The HA, more concerned about the consequences of their actions in regard to anger, tended to find one "safe" way of releasing their tensions and to rely on it rather exclusively. Thus, they were more blocked in selecting appropriate modes of expression of anger depending on the situation.

DISCUSSION

There are three outstanding features of the results. The first is that not one of the findings obtained presented a more favorable picture for the HA children. Each result described the LA as more effective, having fewer and less intense conflicts and better preparation and ability to face and solve the problems of emotional development. Actually many questions did not yield significant results—this for a variety of reasons, such as parental defensiveness, phrasing of questions, and inclusion of many exploratory questions, which were not based on specific predictions or hypotheses. It might be maintained that the fathers of the LA being away overnight more could be regarded as favoring the HA subjects. It is true that the LA children experienced more departures of their fathers which could evoke greater anxiety in those children. However, in all cases these subjects experienced the "safe return" of their fathers, and it is easily conceivable that this was the crucial aspect of the fathers' being away overnight. From that viewpoint the LA children would have had greater opportunity to develop a feeling of security that the departing fathers would return safely.

The second noteworthy feature of these results is not only did the findings favor the LA but the evidence that the mothers of the HA were more defensive, less verbal, and more dependent on the interviewer suggests that further significant findings favoring the LA children were masked or covered up as a result of the greater defensiveness of the HA mothers. Without the differential tendency to censor, the near-significant findings presented might have attained more reliable levels of significance and other analyses of the interview data might have yielded meaningful differences.

The third feature of the findings presented here is that only two of the results were significant only for girls, while significant results were obtained repeatedly for the boys but not for the girls. This is even more striking when it is noted that the difference in anxiety score was greater for the girls than for the boys. The HA girls had higher anxiety scores than HA boys, while LA boys' anxiety scores were only slightly lower than those of the LA girls. In spite of this, more significant differences were found between LA and HA boys than were found for the girls. This could suggest that the anxiety scales are actually "boys' scales" or sex-biased in that they measure

CHILD DEVELOPMENT

the anxieties which can be interfering in boys. This is supported by the finding that the HA boys' school marks were significantly lower than those of the LA boys, whereas the finding for the girls in this respect was one of chance difference. On the other hand, there was the finding, equally significant for girls as well as boys, that the interviewer's write-ups were rated favorably for the LA and unfavorably for the HA children. This finding indicates that the LA girls and boys were seen by the interviewer and an independent judge as more effective and having more favorable experiences in their emotional growth and interpersonal relationships. This result suggests that the TA and GA scales are valid measures of anxiety for both boys and girls.

In evaluating the differential results for boys and girls, perhaps it should be mentioned that there is often a tendency to regard children as a homogeneous group and to disregard the sex differences because similarities are found and broader generalizations *may seem* more desirable than more specific ones. Certainly, there is simple cultural evidence indicating that boys and girls receive different evaluations, treatment, and training. With regard to anxiety girls are permitted to express anxiety and in fact in many respects are rewarded if they are or behave as if they were anxious. On the other hand, boys are rewarded if they are brave little men without fears. Such a view aids in understanding why girls' anxiety scores are higher than boys' and helps explain why admission of anxiety by girls is not necessarily related to interference with their effectiveness, while significant results are found for the boys. Among girls, for example, admitting the fearing of animals may be acceptable and ego-syntonic, and such fears need not reflect an interfering factor on performance or adjustment. But in boys the same fears are unacceptable or ego-alien, and fear is related to conflicts which interfere with the HA boys' successful adjustment.

Some attention, we feel, should be given to the finding that HA and LA girls, as in the case of boys, were significantly differentiated by the evaluation of the interview write-ups, the LA girls being evaluated more favorably. This is one of the very few times that we have obtained the same findings for boys *and* girls. Before attempting to interpret this finding it is necessary to recall that on the basis of teacher grades the LA girls *as a group* did no better than the HA girls. However, in approximately half the cases the LA girl tended to do significantly better than her paired mate while in the remaining cases the HA girl tended to do significantly better than her paired mate. It is possible that among girls an HA score can reflect a high degree of motivation, or the strength of interfering factors, or a combination of both. While an HA score among boys may reflect the same possibilities, it more often than not reflects the strength of interfering factors. Our observations (3) of these same subjects in the classroom fits in well with this interpretation. In that study we found that more HA than LA girls were described as having a high need for achievement, while the observational picture of the HA boys emphasized their inadequacies rather than

KENNETH S. DAVIDSON

their strength of motivation. If these speculations have merit, they suggest as an important problem for future investigation why two girls, both of whom admit to many anxieties, apparently differ in how they are affected by such anxieties. Or to put it another way: It appears necessary to determine whether girls' admission of much anxiety on a questionnaire is experienced as ego-syntonic or ego-alien. From our various studies it appears that among boys such admissions are most frequently experienced as ego-alien.

REFERENCES

1. DAVIDSON, K. S., SARASON, S. B., LIGHTHALL, F. F., WAITE, R. R., & SARNOFF, I. Differences between mothers' and fathers' ratings of low anxious and high anxious children. *Child Developm.*, 1958, 29, 155-160.
2. FOX, CYNTHIA, DAVIDSON, K. S., LIGHTHALL, F. F., WAITE, R. R., & SARASON, S. B. Human figure drawings of high and low anxious children. *Child Developm.*, 1958, 29, 297-301.
3. SARASON, S. B., DAVIDSON, K. S., LIGHTHALL, F. F., & WAITE, R. R. Classroom observations of high and low anxious children. *Child Developm.*, 1958, 29, 287-295.
4. SARASON, S. B., DAVIDSON, K. S., LIGHTHALL, F. F., & WAITE, R. R. Rorschach behavior and performance of high and low anxious children. *Child Developm.*, 1958, 29, 277-285.
5. SARASON, S. B., DAVIDSON, K. S., LIGHTHALL, F. F., & WAITE, R. R. A test anxiety scale for children. *Child Developm.*, 1958, 29, 105-113.
6. WAITE, R. R., SARASON, S. B., LIGHTHALL, F. F., & DAVIDSON, K. S. A study of anxiety and learning in children. *J. abnorm. soc. Psychol.*, 1958, 57, 267-270.

THE HISTORY OF THE

The first part of the history of the world is the history of the creation of the world and the history of the first man, Adam. The second part is the history of the world from the time of Adam to the time of the birth of Jesus Christ. The third part is the history of the world from the time of the birth of Jesus Christ to the present time.

The first part of the history of the world is the history of the creation of the world and the history of the first man, Adam. The second part is the history of the world from the time of Adam to the time of the birth of Jesus Christ. The third part is the history of the world from the time of the birth of Jesus Christ to the present time.

The first part of the history of the world is the history of the creation of the world and the history of the first man, Adam. The second part is the history of the world from the time of Adam to the time of the birth of Jesus Christ. The third part is the history of the world from the time of the birth of Jesus Christ to the present time.

The first part of the history of the world is the history of the creation of the world and the history of the first man, Adam. The second part is the history of the world from the time of Adam to the time of the birth of Jesus Christ. The third part is the history of the world from the time of the birth of Jesus Christ to the present time.

AN EXPERIMENTAL INVESTIGATION OF PARENT-CHILD ATTITUDES WITH THE PARI SCALES

DONALD K. FREEDHEIM and WALLY REICHENBERG-HACKETT*¹

Duke University

This study originated in an endeavor to find a tool for quantitative evaluation of changes in attitudes and behavior of undergraduate students which qualitatively had been evident. The impetus came from seven years of work with students at Duke University. The Psychology Department maintains a nursery group for three- and four-year-olds as a laboratory for child study. It is utilized in a psychology course concerned with "Adjustment of the Preschool Child." The course combined lectures and discussions with work in the nursery group. Students had continuous contact with the children while they were trained in techniques of observation. Objectively observed and recorded behavior was discussed, analyzed, and interpreted in two or three weekly meetings. In these meetings there prevailed an atmosphere of warm acceptance of each child, combined with an endeavor for psychological understanding. The students analyzed behavior items, discussed and interpreted their meaning, and tied them up with theoretical concepts. Optimum opportunities for child growth and development were reviewed and promoted. During the term students frequently noted the effectiveness of measures introduced to assist the "everyday child with its everyday problems." Development of skill in child management on the part of the students appeared to be one by-product of the course. Beyond that, noticeable changes in attitudes and behavior of the participating students became increasingly evident. In the course of recent years this had been stressed by the nursery group teacher in charge, as well as acknowledged by the residency deans of the college.

An evaluation of this class work has continuously been made on three levels: (a) evaluation by the students, (b) evaluation by the trained nursery teachers in charge, and (c) by objective research. The objective research

* Department of Psychology, Duke University, Durham, North Carolina.

¹ The authors gratefully acknowledge support of this work by a grant from the Duke University Research Council.

CHILD DEVELOPMENT

program seemed particularly difficult and complex, calling for criteria as to what constitutes wholesome attitudes towards children and good mental hygiene. As one means of evaluating the impact of the course on attitudes toward children, the PARI Scale (6) was administered to the students in a before and after design. This parent attitude research instrument (PARI) has been developed in recent years because of the ever widening interest in measuring attitudes toward family life, child-rearing practices, and child behavior. It was hoped that changes in the attitudes tapped by this instrument would be related to actually perceived changes in the nursery setting.

While formerly it was held that direct causal relationships exist between certain specific child-rearing practices (i.e., breast feeding, toilet training, etc.) and adult personality, later studies have indicated that recognition of general attitudinal configurations in parents was more useful in predicting the development of behavior patterns. Orlansky (4) has pioneered in bringing this view to light, while the more recent studies by Mark (3), Harris *et al.* (2), and Shoben (7) have elaborated on this point. These experimenters have pointed to a fairly limited number of general orientations such as "democratic atmosphere" (1) or "overprotectiveness" (3) as exemplifying attitude differences among subjects. Drawing upon the most profitable aspects of past work as well as utilizing the present encouraging trends along this line, Schaefer and Bell (6) have devised a multivariate research instrument in the form of a Likert-type questionnaire. It is made up of 23 scales, each concerned with certain aspects of family life, child-rearing situations and relationships. Items used for the scales resulted from experimental tryouts. The questionnaire attempts to partial out the relevant factors contributing to the configurations found to be important attitude variables. A preliminary standardization accompanies the original report. The present study presents an attempt to test the validity and the usefulness of the scales by administering them to several groups, who, though not parents, carried parental responsibility for handling children and whose approach, behavior, and attitude in regard to children were being critically observed.

PROCEDURE

Four groups were tested: Group A included 16 persons on the rehabilitation staff of a children's cerebral palsy hospital. This group comprised highly trained professional staff, responsible for the 24-hour care of 40 to 50 children. The superiority of their approach to children was documented by the very happy atmosphere within this group. These children, physically handicapped and submitting to a rigorous routine, presented a vivid and impressive picture of happy childhood to the casual observer as well as to the consulting psychologist. The effectiveness of the staff's child-centered approach was further demonstrated by the very real progress of these children in every dimension of development. Group B comprised 16 nurses'

aides at the same hospital. This group included young women in training; they were high school graduates from rural North Carolina homes. They worked with the same children as Group A, yet empirically showed distinct differences in attitudes and behavior toward the children. A term of lectures on child development was introduced to provide factual information as well as to promote better understanding of children's needs. Group C included 48 college students taking a course in preschool child adjustment as described in the introduction, and Group D, 47 college students from a history class who served as a control group.

When administering the PARI, we used the 23 scales furnished in the sample by the original authors (6). We added a sixth item to the five items per scale. This sixth item was taken from the pool of statements furnished in the original paper and ranking high among the correlations presented. In accordance with the Likert method, subjects were asked to mark "strongly agree," "mildly agree," "mildly disagree," or "strongly disagree" on each of the particular statements. Twenty of the scales consisted of statements pertaining to attitudes which were considered pathogenic by the authors. These scales have been shown to discriminate more clearly between groups of parents of normal and maladjusted children (6). The three remaining scales represented "healthy" attitudes and were given for rapport purposes. They were not included in the analyses.

The questionnaires were given each group as a group with instructions to work rapidly. College students were asked to put their names on their papers. The hospital staff and nurses' aides took the questionnaire anonymously. A coding method was used to identify tests and retests. Retests were administered to the nurses' aides after a term of lectures on child development; the college students were retested following their semester of intensive discussion and analyses of experienced child behavior. Controls were similarly retested with no such intervening experience. All retests were given approximately four months after the first test. The questionnaires were scored by assigning scores 4 through 1 to the response categories "strongly agree" to "strongly disagree," respectively. The total sum for the six items was found for each subject, and a mean scale score determined for each group.

Compared were results of the staff group with results of the nurses' aides test, the students' experimental group with the staff group, and with the student control group, the results of the first test with results of the second test.

When comparing results of the test and the retest of the student experimental group, there were available also for comparison the records and progress notes kept by the nursery school teachers. These teachers, working in close contact with the students, kept a running report on behavior and attitudes as evident to them in concrete situations. It was customary to evaluate students during their second week with children and at the end of the term.

CHILD DEVELOPMENT

In addition to over-all ratings of progress or nonprogress, there were singled out and defined three areas which, though interrelated, seemed discernible enough to be rated separately on a five point scale, running from 5, the lowest, to 1, the highest score. The areas under consideration were "Perceptiveness," "Relaxed," and "Management." In designing the rating scale teachers' original notes from previous years were utilized. Their verbatim expressions, frequently given in picturesque phrases, were put into an evaluative sequence.

Under "Perception" was estimated the student's awareness of children's needs, sensitivity to moods, awareness of interpersonal relationships within the group, responsiveness to child behavior. Ratings ran from 5—"bull in the china shop behavior," to 1—"sympathetic approach in accordance with child's needs."

Under "Relaxed" was considered a student's confident and secure presence, as well as his approach to children. Ratings ran from 5—"rigid and overconscientious," or "tense and threatened by children," to 1—"outgoing, relaxed and creative, with child's welfare at heart."

Under "Management" consideration was given to a student's ability to deal directly and effectively with situations as they arose. Management implied initiative, ability to use encouragement for motivation, as well as being capable of setting limits. Each student needed to find a workable balance between permissiveness and restriction in his relations with the children, to keep them interested and happy with constructive activities, and with experiences conducive to each child's development. Such "Management" presupposed a mature personality, capable of wisely providing affection and emotional support, while permitting freedom of experimentation and experience. Ratings ran from 5—"a vegetable, unable to manage children," or "quick to correct, rigid, yet irresponsible himself," to 1—"manages children with feeling and understanding, never tampering with behavior but getting into fundamental causes."

RESULTS

The results are presented graphically in Figure 1. The first column names the individual scale. In the following columns deviation scores are plotted for each scale, taking as standard the mean scores of the experimental group of the original authors. The deviations are expressed in stanine scores. The higher scores indicate greater pathogenic weighting.

It may be noted that, on the whole, staff scores are lower than the scores for the nurses' aides; on several scales they differ significantly in what is considered to be the healthier direction. Scores from the first test of the student experimental group are more similar to scores of the staff group than are the nurses' aides scores. On retest of the student experimental group the shift towards a more healthy attitude is quite evident.

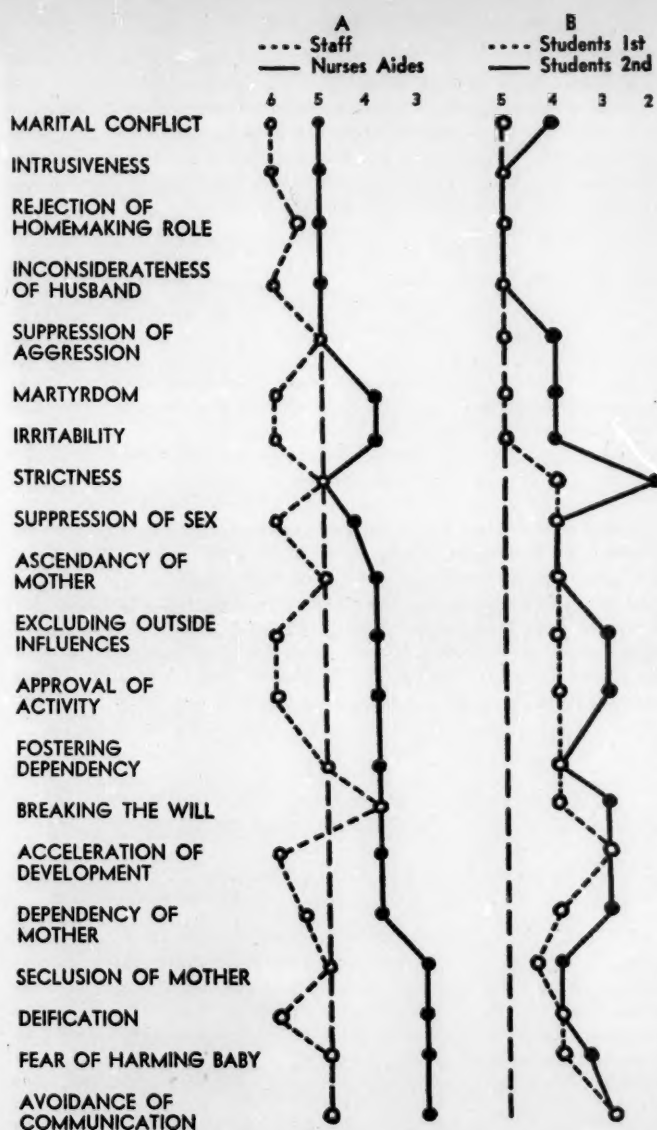


FIGURE 1—Stanine scores for each group on each PARI scale, using as standard the Schaefer and Bell primipara group.

CHILD DEVELOPMENT

TABLE 1

PARI SCALES DIFFERENTIATING SIGNIFICANTLY BETWEEN NURSES' AIDES AND STAFF GROUPS

Scale Name	NURSES N=16		STAFF N=16		t	df	p
	M	SD	M	SD			
Excluding Outside Influences	17.4	3.6	12.3	2.6	4.63	30	.01
Deification	18.0	3.2	12.5	4.2	4.23	30	.01
Acceleration of Development	17.2	3.8	12.3	3.1	4.08	30	.01
Seclusion of the Mother	16.6	2.9	12.3	3.7	3.58	30	.01
Avoidance of Communication	14.5	3.1	11.2	2.8	3.08	30	.01
Martyrdom	13.4	5.0	9.9	3.7	2.59	30	.05
Suppression of Aggression	13.4	3.9	10.8	2.9	2.16	30	.05
Suppression of Sex	12.3	3.6	9.3	4.5	2.14	30	.05
Irritability	16.0	3.5	13.7	2.6	2.09	30	.05

Scores from the first test of the student control group were very similar to scores of the first test of the experimental group. This control group showed no significant differences when retested.

Tables 1 and 2 show the scales which significantly discriminate between groups. In Table 1, nurses' aides versus staff, nurses' aides show significantly higher scores on "Excluding Outside Influences," "Deification," "Acceleration of Development," "Seclusion of the Mother," and "Avoidance of Communication" at the .01 level of significance. At the .05 level, the differentiat-

TABLE 2

PARI SCALES SHOWING SIGNIFICANT SHIFTS UPON RETEST
IN STUDENT GROUP
(N = 48)

Scale Name	TEST		RETEST		t	df	p
	M	SD	M	SD			
Breaking the Will	11.2	2.6	8.7	2.3	5.00	46	.01
Suppression of Sex	8.2	1.9	7.7	1.4	5.00	46	.01
Marital Conflict	17.3	3.1	15.9	3.3	2.80	46	.01
Intrusiveness	12.1	3.3	10.7	3.0	2.80	46	.01
Avoidance of Communication	10.8	2.8	8.4	2.4	2.66	46	.01
Suppression of Aggression	10.9	2.2	9.6	3.3	2.60	46	.05
Approval of Activity	14.3	2.8	13.4	3.4	2.25	46	.05
Deification	13.8	3.5	12.7	3.1	2.20	46	.05
Excluding Outside Influences	12.5	2.3	10.8	2.8	2.12	46	.05

DONALD K. FREEDHEIM and WALLY REICHENBERG-HACKETT

ing scales were "Martyrdom," "Suppression of Aggression," "Suppression of Sex," and "Irritability." Differences on other scales were not significant.

Table 2 shows the significant shifts in scale scores of the student experimental group. "Breaking the Will," "Marital Conflict," "Intrusiveness," and "Avoidance of Communication" are significant at the .01 level. Scales differentiating at the .05 level were "Suppression of Aggression," "Approval of Activity," "Deification," and "Exclusion of Outside Influences." Differences on other scales were not significant.

Changes of nurses' aides' scores at the retest were not significant; however, there was less variation. Scores were closer to the mean, and the group appeared more homogeneous.

The teachers' quantitative evaluation (rating) of students' progress showed a strong trend of linear correlation with PARI scale results. In 45 of the 48 cases, progress or nonprogress moved in the same direction. The four students who gained most, according to behavior records, were among the six students topping the PARI sequence; the same held true at the other end of the scale, for scores of the three poorest students were among the four less healthy scores according to PARI.

A statistical investigation as to which of the scales that differentiate significantly at the retest correspond to the teachers' evaluation of changes in observed behavior is planned for the future.

DISCUSSION

In reviewing the results, it is felt that the differences in scores of staff versus nurses' aides are partly due to the cultural and educational background of the two groups. Very real differences in behavior and orientation towards children are reflected in the scores. The staff was comprised of highly trained professional workers specializing in different fields (physiotherapy, occupational therapy, speech therapy, rehabilitation), all of their education and experience centering on the handling of children. They were competent in promoting their welfare and development. The nurses' aides presented a more inexperienced group, several of them young women in training, others older women starting out on a new career. They brought to the hospital job a variety of attitudes, as well as varying degrees of experience. The lecture course had been instigated because of these noted variations. Four of the scales differentiating between staff and aides ("Acceleration of Development": learned patience; "Seclusion of Mother": social adjustment of mother, acceptance of role; "Martyrdom": parent's sacrifices versus objectified interest; and "Irritability": frustration and short temper with an aura of self-righteousness) reveal the more professional attitude of the staff.

When considering the student experimental group, it can be noted that here are evident shifts in behavior and orientation through training. The increased understanding as evaluated by the teachers is reflected by the attitude assessment.

CHILD DEVELOPMENT

Why the students differed in their second test and the nurses' aides did not differ seems due to the differences in the term's educational program. The nurses' aides were a captive audience; they attended a series of lectures at the end of a busy day, and at times felt the pressure of prolonged hours. They asked questions, but did not seem to be involved beyond a surface level. They requested rules which they could apply.

The students had selected the course of their own accord. Their day by day experiences with children were the subject matter of detailed discussion and analyses. They were eager to understand children; motivation and involvement were high; they enjoyed the children and appreciated whatever behavior became evident. They experienced directly the results of manipulation of emotional components in child behavior. A spontaneous reorganization of attitudes could result; they were ready to rationalize their new beliefs. The nursery group teachers noted the synthesis of cognitive and affective components in the student's progress.

Finally, a word should be added concerning individual scales which differentiate significantly and point up healthier attitudes after training. These scales fall into a pattern, centering around attitudes concerned primarily with immediate interaction with children. It is in scales indicative of a more relaxed and permissive attitude that progress was found in the student group. Among scales not indicative of significant progress are those pertaining more or less directly to domestic and marital situations which at present are not within the range of student experience. No systematic relationship was found among the scales which discriminated within the groups and the five factors brought out in the factor analyses done by Schaefer and Bell (5). Further, the name of a scale cannot be accepted at face value; the question's being asked must be examined. For example, student's progress towards healthier attitudes was noted in regard to "Marital Relations." Studying the questions pertaining to the score, it was found that all of them attempted to reveal the prominence of tensions, for example: "tell people off," "show hurt to achieve change."

PARI was designed by its authors as an instrument to gauge parent-child relationships. It seems that an underlying assumption was made, accepting the emotionally toned husband-wife relationship as the logical or primary source of tensions. The students had other tension-creating experiences; to some of them interacting with children in itself was tension-producing. As they reoriented themselves through mastery of a cognitive structure, they had rewarding insights and experiences; behavior and attitudes changed. They were serene and relaxed, and this change was simultaneously indicated by their score on the scale "Marital Relations."

SUMMARY

Though the results were indicative rather than conclusive because of small numbers, it was found that PARI was sensitive to attitudes of persons carrying parental or quasi-parental responsibilities with young children.

DONALD K. FREEDHEIM and WALLY REICHENBERG-HACKETT

The PARI Scale proved sensitive to differences between two groups, varying in training and background and differing in overt behavior towards children. PARI results reflected shifts in behavior and orientation towards children when student groups, after an educational program, demonstrated increased understanding of child development and child management. The effectiveness of a training course in regard to healthier parent-child attitudes depends on a combination of cognitive information with affective components.

REFERENCES

1. BALDWIN, A. L., KALHORN, T., & BREESE, F. H. Patterns of parent behavior. *Psychol. Monogr.*, 1949, 63, No. 4.
2. HARRIS, D. B., GOUGH, H. G., & MARTIN, W. E. Children's ethnic attitudes: II. Relationship to parental beliefs concerning child rearing. *Child Develpm.*, 1950, 21, 169-181.
3. MARK, J. C. The attitude of mothers of male schizophrenics toward child behavior. *J. abnorm. soc. Psychol.*, 1953, 48, 185-189.
4. ORLANSKY, H. Infant care and personality. *Psychol. Bull.*, 1949, 46, 1-48.
5. SCHAEFER, E. S., & BELL, R. Q. Patterns of attitudes toward child rearing and the family. *J. abnorm. soc. Psychol.*, 1957, 54, 391-395.
6. SCHAEFER, E. S., & BELL, R. Q. Development of a parental attitude research instrument. *Child Develpm.*, 1958, 29, 339-361.
7. SHOEN, E. J., Jr. The assessment of parental attitudes in relation to child adjustment. *Genet. Psychol. Monogr.*, 1949, 39, 101-148.

THE HISTORY OF THE CITY OF BOSTON, FROM THE FIRST SETTLEMENT TO THE PRESENT TIME. BY JOHN H. COLEMAN, ESQ. OF THE BARR.

IN TWO VOLUMES. VOL. II.

BOSTON: PUBLISHED BY J. B. ALLEN, 1827.

THE HISTORY OF THE CITY OF BOSTON, FROM THE FIRST SETTLEMENT TO THE PRESENT TIME. BY JOHN H. COLEMAN, ESQ. OF THE BARR.

IN TWO VOLUMES. VOL. II.

BOSTON: PUBLISHED BY J. B. ALLEN, 1827.

en
of
ca
ru
Th
eff
stu
in
fic

me
th
fo
re
re
pe
th
at
as
th

of
C

THE LATER EFFECTS OF AN EXPERIMENTAL MODIFICATION OF MOTHERING

HARRIET L. RHEINGOLD and NANCY BAYLEY*

National Institutes of Health

An extensive literature in psychology attests to the effect of early experience upon later behavior. For the human infant an important determiner of early experience is maternal care. Some of the dimensions of maternal care thought to be of consequence are amounts and kinds of care, interruptions of care, the number of persons giving care, as well as their attitudes. There is not yet, however, any considerable *experimental* literature on the effects of these variables upon the later behavior of children. The present study reports an attempt to discover the presence, a year later, of a change in behavior brought about in a group of infants by an experimental modification of maternal care (5).

Sixteen children, living in an institution for approximately the first nine months of life, were the original subjects of study. From the sixth through the eighth month of life, half of them, the experimental group, were cared for by one person alone, the experimenter, for 7½ hours a day. They thus received more attentive care than the control subjects who were completely reared under institutional routine; and of course the number of different persons from whom they received care was markedly reduced. As a result the experimental babies became more responsive to the experimenter almost at once, while with time they became more responsive to other persons as well. They did not however do reliably better than the control subjects on the Cattell Infant Intelligence Scale or on tests of postural development and

*Laboratory of Psychology, National Institute of Mental Health, National Institutes of Health, Public Health Service, U. S. Department of Health, Education, and Welfare.

CHILD DEVELOPMENT

cube manipulation. At the conclusion of the study the experimental subjects returned to the full-time care of the institution. Details of the institutional care, of its experimental modification, of the tests used, and of the results may be found in the report referred to above.

One by one, all but one of both the experimental and the control subjects were placed outside the institution—in their own homes, or in adoptive or boarding homes. Approximately a year after the conclusion of the study, the children, then about 18 months of age, were seen again, in an attempt to detect the effects of the earlier treatment. Since the only clear difference between the groups at the time of the study had been an increase in social responsiveness among the experimental babies, it would be here that one would expect a difference, if any, to persist. Still, the possibility existed that differences might appear later as new functions matured. On the other hand, the subsequent, and more recent, experience of several months' duration in different life situations might reduce the chance of finding a difference.

The effects of experimental treatment were sought in two areas of behavior, the social and the intellectual. Would the experimental subjects be more socially responsive, that is, more friendly and outgoing than the control group to two examiners who visited the home? Would the experimental subjects, in addition, be more responsive to the original experimenter than to another person? If not, the variable under test is really their responsiveness to strangers. Second, would the experimental subjects now do better on a test of developmental progress?

It was planned, in addition, to use the retest data to explore the effect of type of home placement, as well as to evaluate the performance of the whole group considered as a sample of institutionalized children.

PROCEDURE

Subjects

Fourteen of the original 16 children were located and tested; one from the experimental group and one from the control group could not be found.¹

The mean age of the experimental group was 19.8 months (range, 17.6–22.1), of the control group, 20.1 months (range, 17.5–21.7). The experimental group had spent an average of 9.2 months in the institution before being placed in homes (range, 4.0–13.6); for the control group the mean time was 10.4 months (range, 6.5–18.1). If the control subject who was still in the institution was omitted from the calculations, the average stay for the control group became 9.2 months (range, 6.5–12.2). In respect, then, to age and to duration of stay in the institution both groups were similar.

The children left the institution at different ages. Two experimental subjects left after only three weeks of treatment. One control subject left

¹ We are grateful to Father Bernard Brogan, Director of the Catholic Home Bureau of Chicago, for his generous cooperation.

HARRIET L. RHEINGOLD and NANCY BAYLEY

in the sixth week of the study, another in the seventh week. All the other subjects stayed at least through the eight weeks of treatment.

The home placements were varied. Three experimental and two control subjects returned to their own homes. With one exception, the own parents of these five subjects were of foreign birth and the homes were marked by poverty. Two of the experimental and four of the control subjects were in adoptive homes which, in general, were superior to the own homes in socioeconomic status. Two experimental subjects were living in boarding homes, pending a release for adoption. And one control subject, a Negro boy, remained in the institution only because a home could not be found for him. Furthermore, there was no difference between the experimental and the control groups in the intellectual stimulation provided by the homes or in the friendliness of the mothers, according to ratings made by the Experimenter and the Examiner after each visit. In type of home placement, therefore, there appeared to be no major difference. Rather, the difference between homes within each group appeared to be larger than any difference between the groups.

The Tests

Each child was seen in his own home. The homes were scattered widely through Chicago, its suburbs and neighboring cities, with one home in another state. Two persons, the original Experimenter and an Examiner, visited the homes together, with one exception: the child who lived out of the state was examined by the Experimenter alone. The Experimenter knew all the children but, of course, had been especially familiar with the experimental subjects. She served *only* as a stimulus person in the social tests. The Examiner had no previous acquaintance with any of the children and did not know which had been the experimental subjects. She also served as a stimulus person in the social tests, but it was she alone who recorded the children's responses to both the Experimenter and herself, and who administered the test of developmental progress.

The social test resembled those reported in the first study, but was made more suitable for older children. It was composed of three parts, each of which set up a rather natural situation between adult and child, with an easy transition between the parts. In the first part, the responses to the stimulus person in the first few minutes after her entrance into the home were recorded. During this time the stimulus person did not talk to or approach the child but sat at some distance from him and talked to the mother, occasionally smiling at the child. The Examiner recorded the child's responses to whichever stimulus person happened first to engage his attention, then to the other person. At an appropriate moment one of the persons smiled and spoke warmly to the child, saying, "Hi (child's name) come to me," accompanying her words by stretching out her arms in invitation. This constituted the second situation. In the third situation, the stimulus person actually approached the child, smiling, talking, and gesturing as

CHILD DEVELOPMENT

in the second situation. After the child's responses had been recorded, the other stimulus person presented herself to the child similarly. The order of stimulus persons was determined by the convenience of the moment: whoever was closer to the child or was receiving more glances was the first stimulus person.

The child's responses were recorded on a checklist under these categories: *positive facial expression*, which included seven items of behavior ranging from "stares with expression" to "laughs"; *physical approach* with nine items ranging from "shows toy" through "makes physical contact" to "makes social overtures while in the stimulus person's lap"; *vocalizations* for which a child received a score of one for each part of the test in which he vocalized, whether he said discrete sounds, jargon, or words; *negative facial expression*, which included eight items ranging from "a fleeting sober expression" to "cries"; *physical retreat* with six items ranging from "hangs head" to "leaves room"; and *response to mother* (during the social test period) which included a series of six items, from "turns toward mother" to "stays in contact with mother."

Within each category, items of behavior were thus arranged in what seemed a reasonable progression in terms of duration or amplitude of response. Each item within a category was arbitrarily assigned a value of one. Because the items were arranged in ascending order, the score for any item was one plus the value of all other items below it in that category. The scores for the categories of positive facial expression, physical approach, and vocalizations were summed to yield a measure of *positive social responsiveness*. Similarly, the sum of both negative categories gave a measure of *negative social responsiveness*. The sum of these two measures was the measure of *total social responsiveness*. The category of "response to mother" was calculated separately and not included in the other measures.

After the social tests, the Cattell Infant Intelligence Scale (2) was administered by the Examiner, with the Experimenter *not* present. Lastly, the number of words in the child's vocabulary was calculated from his performance on the language items of the Cattell and from the mother's report.

RESULTS

The Effect of Treatment

Table 1 shows that both the experimental and the control subjects responded similarly to the Experimenter and to the Examiner. The close agreement of all means, and of the ranges, is apparent in the part, as well as in the total, scores. The only difference of any size between the two stimulus persons appeared in the experimental group's response to mother. But since only one subject of the seven gave a response to the mother when the Experimenter was the stimulus person, and only three subjects of the seven, when the Examiner was the stimulus person, this difference, as all the others, was not statistically significant. From the results we conclude

TABLE 1
MEANS AND RANGES OF THE SOCIAL TEST

<i>Subjects</i>	<i>Experimenter</i>		<i>Examiner</i>		<i>Combined Score</i>
	<i>Mean</i>	<i>Range</i>	<i>Mean</i>	<i>Range</i>	<i>Mean</i>
<i>Experimental Group*</i>					
Total Social Responsiveness	32.1	27-39	30.9	27-38	31.6
Positive	17.4	2-30	16.0	2-37	16.7
Negative	14.7	1-37	14.7	3-29	14.7
Response to Mother	2.3	0-16	5.7	0-19	4.0
<i>Control Group†</i>					
Total Social Responsiveness	28.0	14-39	28.3	22-44	28.4
Positive	19.8	5-32	20.2	4-37	20.1
Negative	8.2	3-12	8.2	2-18	8.0
Response to Mother	4.5	0-11	4.8	0-10	5.4

* *N* is 7.

† *N* is 6 for responses to Experimenter and to Examiner, but 7 for Combined Score. See text for explanation.

that the experimental subjects did not appear to remember the Experimenter.

Furthermore, since the experimental and the control groups gave similar scores to both persons, it was assumed that they were of approximately equal stimulating value. Therefore, a combined score for each subject (the average of a subject's responses to both stimulus persons) was used in the analyses which follow. This procedure made it possible to add to the control group the subject who was seen by the Experimenter alone. If every other subject responded similarly to both stimulus persons, it may be assumed that this subject would too. (It will be seen in Table 1 that the addition of this subject to the control group made the combined means slightly different from the separate means.)

The combined scores showed that the experimental subjects were more responsive to both persons than the control subjects, but the difference was not statistically significant. The part scores, further, revealed that the control group gave more positive responses, the experimental group, more negative responses. Again, the differences were not statistically reliable. Moreover, inspection of the data revealed that the negative responses of only two of the seven experimental subjects were responsible for the difference between the groups. The findings therefore do not warrant the conclusion that the experimental subjects were either more or less responsive to the stimulus persons, positively or negatively.

Because some of the subjects made no response to their mothers during the social tests, the means for this category of behavior were not subjected

CHILD DEVELOPMENT

to test. Only three of the seven experimental subjects and five of the control subjects made some contact with the mother during social stimulation by one or the other of the stimulus persons, a difference which permits no conclusive statement of difference.

Although vocalizations had been included in the measure of positive social responsiveness (as explained above), a measure which did not differentiate the groups, they were also analyzed separately. Inspection showed that five of the seven experimental subjects vocalized to one or the other of the stimulus persons but only one of the control subjects did. The difference was significant by the Fisher exact probability test at $p = .051$ (one-sided), a finding in agreement with the original study in which, at the end of the experimental treatment, the experimental subjects also vocalized more than the control subjects.

On the Cattell Infant Intelligence Scale the mean IQ for the experimental group was 97.4 (range, 82-110); for the control group it was 95.4 (range, 83-122). More attentive care given during a limited period in the first year of life therefore appeared to produce no difference in IQ on retest a year later.

The experimental subjects had a larger spoken vocabulary than the control subjects (17.9 and 13.7 words), but the difference was again not statistically significant.

The Effect of Home Placement

It early became clear that the adoptive homes were of a higher socioeconomic level than the own homes, and therefore it seemed desirable to look for differences in the performance of the children in these two types of home placement. The adoptive homes were also ranked higher than the own homes by the investigators on the basis of the friendliness of the mother during the visit and of the intellectual stimulation the home seemed to offer the child.

On the social test the children in adoptive homes gave more positive responses than those in own homes; the means were 21.6 and 15.6, respectively, but the difference was not statistically significant. It should be noted, however, that one subject in a boarding home and the subject still in the institution made higher positive scores than the mean of the adoptive home group.

Similarly, the mean IQ of the children living in adoptive homes was higher (98.8) than that of those living in own homes (95.4), but the difference was not reliable. The two children living in boarding homes had IQs of 95 and 102. And, while the child still in the institution obtained an IQ of only 83, two children in own homes had lower IQs, one of 79 and one of 82, and one child in an adoptive home had an IQ of 84.

Finally, the children in adoptive homes had a larger vocabulary than the children in own homes (means were 18.6 and 13.4, respectively), although again the difference was not significant.

In summary, there was no reliable evidence that the children in adoptive homes were more socially responsive or more developmentally advanced than those in own homes.

The Group as a Whole

We may now evaluate the performance of the group as a whole ($N = 14$), representing as it does a sample of children who spent approximately nine months of the first year of life in the care of an institution and who then experienced a major change in life situation.

In general, the group was marked by a friendliness which seemed warm and genuine. Eleven of the 14 Ss not only approached the stimulus persons but also allowed themselves to be picked up and held. Only two subjects, both boys, presented a different social response: they clung to their mothers and cried when the stimulus persons approached them. No comparable data are available for children who have lived all their lives in own homes, but in preliminary testing of the social test on three such children not one approached the examiners. Instead, they looked at the examiners from behind their mothers' skirts and retreated whenever the examiners moved in their direction.

On the Cattell Infant Intelligence Scale the mean IQ of the group was 96.4. At six months of age the mean IQ for these 14 children was 93.8; at eight months it was 94.3. They continue therefore to score in the normal range. Furthermore, the mean number of words in their vocabulary was 15.5, which compares favorably with Gesell's (3) norms of 10 words at 18 months and 20 words at 21 months. Certainly, the group showed no sign of mental dullness or of language retardation.

No child, furthermore, showed the marked apathy or attention-seeking behavior believed by some to characterize the behavior of children reared in institutions. Differences there were, to be sure, between the children, but none seemed to depart markedly from the normal in temperament or personality. In fact, several of the mothers spontaneously commented upon how easy these children were to handle in comparison with their other children. They mentioned, specifically, their good eating and sleeping habits and their ability to amuse themselves.

DISCUSSION

The discussion will take up three separate points: (a) the effect of the experimental treatment, (b) the effect of own home versus adoptive home placement, and (c) the characteristics of the whole group considered as a sample of institutionalized children.

On the basis of the changes in social behavior produced at the time of treatment, one might have expected that the experimental subjects on retest would have been more responsive to the Experimenter than to the

CHILD DEVELOPMENT

Examiner. Instead, no reliable difference was found in their responses to either person. The Experimenter was not remembered. Further, we did not find, except in the vocalizing of the children, any evidence that the experimental subjects were more responsive than the control subjects. It seems, therefore, that the experiences provided by the more attentive mothering were not great enough to maintain the experimentally produced difference over a year's time, except in one class of behavior.

The findings give rise to several speculations. First, it is possible that the verbal behavior of young children is more sensitive to changes in the environment than are other classes of behavior. In this connection, the responsiveness of vocalizations to conditioning in the three-month-old infant has already been demonstrated (6). Second, differences between the experimental and control groups may well have existed but in some untested area of behavior. Third, the expected (or some other) differences may make their appearance in the future in some more complex behavior incorporating the experiences of treatment. Finally, serious limitations to the study were imposed by the small number of subjects and by the diversity of home placements within each group. Differences would have to be very large indeed to surmount these limitations.

That no difference was found between the experimental and control groups in developmental status is not surprising, considering that no difference was found at the end of treatment. Some of the speculations about the course of social responsiveness may apply here, too.

We turn now to a consideration of the effect of home placement. The adoptive homes in general were of a higher socioeconomic level, the mothers were more sociable, and the homes were judged to offer more intellectual stimulation. For these reasons we would have expected the children in adoptive homes to be more socially responsive and more advanced in developmental status. But significant differences were not found. Possible explanations are that the differences between the two groups of home may have been not as great as they seemed, or that the number of cases was too small.

Lastly, the characteristics of the group as a whole may be assessed for the effects of a life experience usually thought of as deprived. All the children had been cared for in an institution for the first half of their lives, all but one had experienced a major "separation" in going from one life situation to another, and, furthermore, three children were now living in depressed socioeconomic environments, two were in boarding homes, and one was still in the institution. Yet, as a group, the children were healthy, of normal intelligence, and they appeared to be making a satisfactory adjustment. In addition, they seemed to be more friendly to strangers than children who have lived all their lives in own homes and, according to mothers' reports, were more adaptable than their other children. In no way, then, did they resemble the emotionally disturbed and mentally retarded children described in studies of the effect of institutional or hospital life or

of separation from the mother.² They did not show apathy or the inability to form relationships or make excessive bids for attention. Even earlier, at the beginning of the study when the infants were still in the institution, they were physically robust, mentally alert, and socially responsive.

It is true that in kind and duration of experience they resemble exactly no other group of children reported in the literature. There is a tendency among workers, however, to lump together studies of children who actually differ in age and experience and to generalize from them to all children who have experiences which may be similar in only one of many possible respects. It is to be hoped that as more prospective (in contrast to retrospective) studies are carried out, the dimensions of deprivation and of its effects can be clarified. Certainly, we may expect to find that the effects will depend upon the age of the child, the nature and duration of the deprivation, and the experiences prior to and subsequent to it (1). The present study of the effects of early experience, limited as it is, emphasizes the need for more precise measurement both of deprivation and of its effects.

SUMMARY

The present study reports an attempt to discover the presence, a year later, of a change in behavior brought about in a group of infants by an experimental modification of maternal care.

Sixteen babies, living in an institution for approximately the first nine months of life, were the original subjects of study. Half of them, the experimental subjects, received more attentive care by one person, the Experimenter, from the sixth through the eighth month of life. As a result they became more socially responsive than the control group who were cared for under the usual institutional routine. They did not, however, do better upon tests of developmental progress.

Subsequently all but one of the children were placed in homes. A year later, when the children were about 19 months old, 14 of the original 16 subjects were located, and tested for their social responsiveness and developmental progress.

The results did not reveal any statistically significant differences between the experimental and the control groups except that more of the experimental subjects vocalized during the social tests. It is concluded therefore that the experience provided by the more attentive mothering, while great enough to produce a difference at the time of study, was not great enough to maintain that difference over time, except in one class of behavior. It is possible that the verbal behavior of young children is more sensitive to changes in the environment than are other classes of behavior.

No statistically significant differences in social responsiveness and developmental status were found between children living in own homes and

² Glaser and Eisenberg (4) present a recent review of studies on maternal deprivation.

CHILD DEVELOPMENT

in adoptive homes, although the adoptive homes were of higher socioeconomic status.

Finally, the group as a whole was friendly, of normal intelligence, and apparently was making a satisfactory adjustment. They did not resemble the emotionally disturbed and mentally retarded children described in studies of the effects of institutional life or of separation from the mother.

REFERENCES

1. AINSWORTH, MARY D., & BOWLBY, J. Research strategy in the study of mother-child separation. *Courier*, 1954, 4, 105-131.
2. CATTELL, PSYCHE. *The measurement of intelligence of infants and young children*. New York: Psychol. Corporation, 1940.
3. GESELL, A., & AMATRUDA, CATHERINE S. *Developmental diagnosis*. New York: Hoeber, 1941.
4. GLASER, K., & EISENBERG, L. Maternal deprivation. *Pediatrics*, 1956, 18, 626-642.
5. RHEINGOLD, HARRIET L. The modification of social responsiveness in institutional babies. *Monogr. Soc. Res. Child Develpm.*, 1956, 21, No. 2. (Serial No. 63)
6. RHEINGOLD, HARRIET L., GEWIRTZ, J. L., & ROSS, HELEN W. Social conditioning of vocalizations in the infant. *J. comp. physiol. Psychol.*, 1959, 52, 68-73.

THE MEASUREMENT OF MASCULINITY AND FEMININITY IN CHILDREN¹

B. G. ROSENBERG and B. SUTTON-SMITH*

Bowling Green State University

In 1936 Terman and Miles published an attitude-interest-analysis test which was constructed so as to yield a measure of masculine-feminine differences (10). This test originated in their discovery that selected children's games and play activities could be used to discriminate between the sexes (9). A recent paper by the authors has shown that the games which Terman and Miles found would discriminate between the sexes in 1926 do not always discriminate between them today. There has been considerable change in children's games during the intervening years (6). The general agreement that substantial masculine-feminine differences in interests do exist (2, 5, 7), the predictive effectiveness of this type of masculine-feminine scale (10, p. 16), the scale's ability to yield information on other important psychological variables (10), and its comprehensibility by children suggest the value of a revised scale for use with children today. This paper is concerned with the construction and validation of such a scale.

PROCEDURE

A list of 181 games was administered to a sample of 183 children in the fourth, fifth, and sixth grades.² The children were instructed to check like or dislike for only those items they recognized. A final list of 115 items

* Department of Psychology, Bowling Green State University, Bowling Green, Ohio.

¹ This study was facilitated by a grant from the Scholarly Advancement Committee, Bowling Green State University.

² A mimeographed copy of the game check list can be obtained from the authors upon request.

CHILD DEVELOPMENT

TABLE I

ITEMS AND WEIGHTS INCLUDED IN THE MASCULINITY AND FEMININITY SCALES *

MASCULINITY (56 possible)		FEMININITY (116 possible)	
Item No.	Game	Item No.	Game
W E I G H T 3			
9.	soldiers	1.	dolls
13.	cops & robbers	2.	dressing up
18.	spacemen	3.	houses
64.	bows and arrows	4.	store
65.	throw snowballs	5.	school
113.	darts	23.	red rover
128.	wrestling	27.	fox and geese
131.	football	30.	stoop tag
135.	boxing	34.	ring around the rosey
139.	shooting	35.	London Bridge
164.	use tools	36.	farmer in the dell
167.	hunting	38.	in & out the window
172.	make model airplanes	41.	drop the handkerchief
173.	toy trains	42.	huckle buckle beanstalk
W E I G H T 2			
8.	bandits	44.	mulberry bush
14.	cars	45.	jump rope
53.	marbles	47.	hopscotch
179.	build forts	49.	jacks
W E I G H T 1			
10.	cowboys	54.	follow the leader
67.	wall dodgeball	56.	crack the whip
68.	king of the mountain	59.	leap frog
127.	bowling	71.	Simon says
130.	soccer	76.	name that tune
137.	racing	78.	musical chairs
W E I G H T 2			
		80.	hide the thimble
		87.	Mother, May I?
		91.	statues
		153.	scrapbook making
		156.	dance
		159.	sewing
		161.	cooking
		174.	see saw
		176.	cartwheels
		7.	doctors
		21.	blind man's buff
		74.	I've got a secret
		110.	puzzles
		111.	clue
		114.	pick up sticks
		180.	build snowmen
		73.	spin the bottle
		82.	I spy
		181.	clay modeling

* Only "likes" are scored.

was obtained to which at least 50 per cent of the children had responded. An analysis was conducted on the basis of sex preference. This analysis yielded 67 items useful in differentiating boys and girls. A scale was developed which assigned simple weights corresponding to the level of significance obtained. Thus, items which differentiated boys from girls at the .01 level were weighted 3, items which differentiated $>.01 < .05$ received a weight of 2, and items which differentiated $>.05 < .10$ were assigned a weight of 1. The same technique was used for the items differentiating girls from boys. The scales were then administered to two subsequent samples in order to determine their validity.

TABLE 2

DISTRIBUTION OF SCORES ON THE MASCULINITY AND FEMININITY
SCALES FOR THE CRITERION SAMPLE

MASCULINITY			FEMININITY		
Scores	Males (N=91)	Females (N=92)	Scores	Males (N=91)	Females (N=92)
54-56	9		113-117		2
51-56	10	1	108-112		4
48-50	13	2	103-107		6
45-47	10	3	98-102		4
42-44	11	2	93-97	1	10
39-41	8	2	88-92	1	11
36-38	10	3	83-87	4	10
33-35	6	6	78-82	0	10
30-32	2	2	73-77	2	10
27-29	5	3	68-72	3	4
24-26	3	5	63-67	4	3
21-23	2	9	58-62	5	6
18-20	0	5	53-57	3	1
15-17	0	10	48-52	9	0
12-14	1	16	43-47	9	2
9-11	0	10	38-42	6	3
6-8	1	7	33-37	9	2
3-5		5	28-32	5	3
0-2		1	23-27	10	0
			18-22	5	1
			13-17	8	
			8-12	5	
			3-7	2	
			0-2		
Mean	41.92	20.32		40.50	79.62
SD	9.49	12.37		21.63	21.15

CHILD DEVELOPMENT

RESULTS

The items and weights of the masculinity and femininity scales are presented in Table 1. It is possible to obtain a score of 56 on the masculinity scale and a score of 116 on the femininity scale. As can be seen, fewer items differentiated boys from girls than girls from boys. Treating each group of items as a separate scale, an analysis for masculinity and femininity was conducted on the criterion sample of boys and girls comprising the fourth, fifth, and sixth grades at a local elementary school. Table 2 presents the results of this analysis seeking to determine to what extent the "scales" could differentiate the criterion sample. The separation of males and females is satisfactory. Only 9 per cent of the females fall above the mean of the males on the masculinity scale, and only 2 per cent of the males fall below the female mean score on this scale. The results are similar for the femininity scale, as 10 per cent of the females fall below the mean of the males, and only 7 per cent of the males fall above the mean of the females. The differences between males and females on the masculinity and femininity scales are highly significant ($p = .001$).

In an attempt to validate these findings, the scales were then administered to 73 boys and 92 girls in the fourth, fifth, and sixth grades at another elementary school in the same city. For the masculinity scale, the boys had a $M = 42.89$ and $SD = 9.78$, while the girls had a $M = 18.42$ and $SD = 11.04$. The difference is significant beyond the .001 level. A second cross-validation of the masculinity scale was conducted at an elementary school 60 miles away.³ The results are highly similar for children in the fourth, fifth and sixth grades: for 199 boys, the $M = 39.81$, $SD = 11.01$; for 207 girls, the $M = 19.42$, $SD = 13.05$. The difference between means is 20.384, significant beyond the .001 level.

The validation of the femininity scale yielded similar results: for 75 boys,⁴ the $M = 44.77$, $SD = 23.54$. For 90 girls, the $M = 76.00$, $SD = 22.43$. Again, the results are similar to those findings with the criterion group on the femininity scale. For the second validation group, the 199 boys had a $M = 41.56$ and $SD = 22.02$. For 205 girls, the $M = 74.05$, $SD = 23.01$. The results attain the .001 level of significance.

Table 3 presents the distributions of scores for the two validation samples on the masculinity scale. It can be noted that for the first validation sample on the masculinity scale, only 2 per cent of the females fall at or above the mean of the males, while 4 per cent of the males fall below the mean of the females. For the second validation sample, 9 per cent of the females fall above the mean score of the males, while 5 per cent of the males fall below the mean score of the females.

³ The authors are indebted to Dr. B. Rabin for his assistance in obtaining this sample.

⁴ The slight differences in N s on the masculinity and femininity scales are due to incomplete protocols which were eliminated.

TABLE 3
DISTRIBUTION OF SCORES ON THE MASCULINITY SCALE FOR
TWO VALIDATION SAMPLES

Scores	VALIDATION 1		VALIDATION 2	
	Males (N=73)	Females (N=92)	Males (N=199)	Females (N=207)
54-56	7		12	
51-53	8		24	1
48-50	14		24	3
45-47	9	1	17	8
42-44	11	2	22	4
39-41	6	0	21	6
36-38	5	4	19	11
33-35	3	4	15	4
30-32	5	8	11	10
27-29	0	7	8	13
24-26	1	8	7	10
21-23	1	4	5	14
18-20	1	6	5	18
15-17	1	8	4	19
12-14	0	9	2	13
9-11	0	13	1	20
6-8	1	4	1	25
3-5		10	1	20
0-2		4		8
Mean	42.89	18.42	39.81	19.42
SD	9.78	11.04	11.01	13.05

The distributions of scores for the two validation samples on the femininity scale are presented in Table 4. For the first validation sample on the femininity scale, 8 per cent of the males fall above the mean score of the females, while only 7 per cent of the females fall below the mean of the males on this scale. For the second validation sample, 8 per cent of the males fall above the mean score of the females, while 6 per cent of the females fall below the mean of the males.

The separation on both cross-validation samples appears acceptable. The *t* test of the differences between the means for the masculinity scale yields values of 14.802 and 16.598, both significant well beyond the .001 level. The *t* test values for the femininity scale are 8.675 and 14.549, both of similar significance.

CHILD DEVELOPMENT

TABLE 4
DISTRIBUTION OF SCORES ON THE FEMININITY SCALE FOR
TWO VALIDATION SAMPLES

Scores	VALIDATION 1		VALIDATION 2	
	Males (N=75)	Females (N=90)	Males (N=199)	Females (N=205)
113-117		1	1	1
108-112		5	0	3
103-107	1	5	1	7
98-102	1	6	0	14
93-97	1	3	1	14
88-92	1	11	4	20
83-87	1	10	2	17
78-82	1	7	6	17
73-77	1	7	2	21
68-72	6	3	8	23
63-67	7	7	11	18
58-62	4	4	10	11
53-57	1	6	17	7
48-52	6	7	9	8
43-47	8	3	19	12
38-42	7	0	15	2
33-37	4	0	16	2
28-32	6	1	17	3
23-27	6	2	23	0
18-22	1	1	8	0
13-17	4	0	11	2
8-12	4	1	10	1
3-7	4		6	2
0-2			2	
Mean	44.77	76.00	41.56	74.05
SD	23.54	22.43	22.02	23.01

DISCUSSION

The finding that a check list of play activities and games will discriminate effectively between the sexes, as well as the discovery that only a small percentage of each sex fall beyond the mean of the other sex, is essentially similar to Terman's findings in 1926 (9, p. 13).

It is of interest to note the nature of the items which discriminate between the sexes in the 8- to 12-year range. Though the items were derived and validated empirically, they are much like those items which might have yielded from *a priori* judgments. The distinctive boys' games involve

forceful physical contact, the dramatization of conflict between male roles, the propulsion of objects over space, and complex team organizations. The distinctive girls' games involve the dramatization of men or women in static activity, the verbal elements of singing or chanting, ritualistic non-competitive actions, choral activities, rhythmic activities, and games in which a central person player has a focal role (8). In general, these differences are consistent with earlier findings (3, 4, 5), and with theoretical conceptions of sex differences in children (1). There is, however, a tendency for girls to show more preference for active, outdoor, competitive and team games than was the case in earlier investigations.

It is suggested that one or both of these scales based on a check list of gaming activities is of practical and theoretical value in the detection of variation in sex role identifications in children.

SUMMARY

The present paper describes the development of a pair of scales measuring masculinity and femininity in children. The scales were empirically derived using game items in a check list, and resulting in an easily administered, comprehensible, nonthreatening, and valid set of items. The scales were cross-validated on two independent samples, and yielded acceptable levels of discrimination.

REFERENCES

1. ERICKSON, E. H. Sex differences in the play configurations of American pre-adolescents. In M. Mead & M. Wolfenstein (Eds.), *Childhood in contemporary cultures*. Chicago: Univer. of Chicago Press, 1955. Pp. 324-341.
2. FORD, C. F., & TYLER, L. E. A factor analysis of Terman and Miles' M-F test. *J. appl. Psychol.*, 1952, 36, 251-253.
3. FOSTER, J. C. Play activities of children in the first six grades. *Child Develpm.*, 1930, 1, 248-254.
4. HILDRETH, C. Adolescent interests and abilities. *J. gen. Psychol.*, 1945, 43, 65-93.
5. LEHMAN, H. C., & WITTY, P. A. *The psychology of play activities*. New York: Barnes, 1927.
6. ROSENBERG, B. G., & SUTTON-SMITH, B. A revised conception of masculine-feminine differences in play activities. *J. gen. Psychol.*, in press.
7. STRONG, E. K. *Vocational interests of men and women*. Palo Alto: Stanford Univer. Press, 1943.

CHILD DEVELOPMENT

8. SUTTON-SMITH, B. A formal analysis of game meaning. *Western Folklore*, 1959, 18, 13-24.
9. TERMAN, L. *Genetic studies of genius*. Vol. I. Palo Alto: Stanford Univer. Press, 1926.
10. TERMAN, L. M. & MILES, C. C. *Sex and personality*. New York: McGraw-Hill, 1936.

ORDERING OF PHENOMENON IN A PAIRED COMPARISONS PROCEDURE

EDWARD CLIFFORD*^{1,2}

Child Research Council, University of Colorado Medical School

A small number of studies have utilized the paired comparisons procedure with nursery school children. Koch (4) and Lippitt (5) used the procedure to study popularity in preschool children, while Clifford and Wischner (1), and more recently Guyette (3), used the method to study children's judgments of punishments. All of these investigations report relatively low reliabilities, measured as the proportion of times a child agrees with himself from session to session. While reliabilities necessarily vary, the reported reliability for all studies is about .43.

When one investigates the data of a particular study (1), a curious phenomenon is seen. The method of paired comparisons assumes a necessary logical relationship between the items to be scaled. Thus, if A is preferred to B and B is preferred to C, it should follow that A is also preferred to C. With preschool children, however, they may indicate, despite logic, that C is preferred to A. In addition to these inversions of logic within a particular session, the judgments do not seem to be stable over a period of time. The lack of stability may be due to the types of items ordered in this procedure. It is possible that where there is a "correct" or predetermined order to the items, both logical consistency and reliability between sessions could be increased. Thus, for example, if items were prescaled along a physical dimension, and if the separation between items were great enough, greater reliability could be expected. Data have been obtained on popularity and punishments, and these are essentially social values about which it may be impossible for children at this age level to make stable judgments. Thus, a study encompassing both types of judgments may enable one to trace the growth of the awareness of logical relationships. A developmental study is needed to determine where, along the age continuum, there is exhibited stability of choice and the incorporation into thinking of the logical relationships needed for adequate use of the paired comparisons technique.

* Child Research Council, University of Colorado Medical School, Denver.

¹ The author wishes to express his appreciation for the cooperation received from the University of Massachusetts Nursery School, Mrs. Emily Thies, director, and to Mrs. Miriam Clifford and Mrs. Florence Leiman for assistance in collecting the data.

² This study was conducted at the University of Massachusetts.

CHILD DEVELOPMENT

EXPERIMENTAL DESIGN

Eighteen children (10 girls, eight boys), ranging in age from 3 years, 4 months to 4 years, 9 months, enrolled in the University of Massachusetts Nursery School, served as subjects. All children were seen individually in a test (Session I) - retest (Session II) situation.

Four scales were constructed for use in a paired comparisons procedure. Scale O consisted of the names of eight physical objects ranging in size from a pin to a house. This scale has an inherent order and should obtain the highest reliability. Scale F consisted of the names of eight fruit. This scale can possibly be equated to a popularity scale in the sense that it demanded judgments of like-dislike. Scale P consisted of eight items of punishment taken directly from a study by Clifford and Wischner (1). The items on this scale represent social values. Scale R consisted of eight reward items which were designed to parallel the categorization of the punishment items and also represent social values.

Using the paired comparisons procedure, each scale contained a total of 28 items. Each item was printed on a card. Before presentation the experimenter shuffled the cards, assuring a random presentation of the order of items for each subject. In addition, two sets of cards were prepared. Set A for each scale consisted of the paired 28 items. Set B consisted of the identical items, with the position of the items reversed. This was done to eliminate any effect of position preference. Half the children were given Set A in Session I, half Set B. In Session II the child was given the set he had not been given in the first session. Scales O, F, P, R were given in order, and for each scale there was an introductory story read to the child. All items were given verbally.

The introductory story and items used in each of the four scales are given below:

SCALE O

Introductory Story

I wish to play a game with you. The game is called "WHICH IS BIGGER?" I am going to name two things and I want you to tell me which is bigger. Listen!

Which is bigger, an elephant or a mouse?

If the child responds correctly say, "That's right," and proceed to the first pair.

If the response is incorrect, point out the distinction and try this one: "Which is bigger, your thumb or your head?"

Then proceed to the first pair.

Items

Pin
Spoon
Dish
Pot
Chair
Table
Car
House

EDWARD CLIFFORD

SCALE F

Introductory Story

I want to play a game about fruit. You know what fruit are, don't you? Let's try and name some fruit. (Name a variety of fruit with the child.) Now I'm going to name two fruit and you tell me which fruit you like better.

Which do you like better or

Items

Bananas
Apples
Pears
Oranges
Grapes
Grapefruit
Peaches
Cherries

SCALE P

Introductory Story

You know that sometimes when children are naughty (bad) their mother might punish them. (Pause).

Now there is this boy (girl) Mike (Mary) and his (her) mother. He (she) did something naughty and his (her) mother decided to punish him (her).

Which would be worse for Mike (Mary)?

Would it be worse

Items

If his (her) mother said, "I'll send the bogeyman after you."

If his (her) mother said, "I'll drive off in the car and leave you all alone in the house."

If his (her) mother took away his (her) favorite toy (dessert).

If his (her) mother spanked him (her).

If his (her) mother scolded him (her).

If his (her) mother said, "You are just like a little baby."

If his (her) mother sent him (her) up to his (her) room (or to bed).

If his (her) mother said, "I won't love you any more."

Category

Supernatural Force

Desertion

Deprivation

Spanking

Scolding

Shaming

Isolation

Withdrawal of Love

SCALE R

Introductory Story

You know that sometimes when children are good, their mother might do something nice for them. (Pause).

Now there is this boy (girl) Mike (Mary) and his (her) mother. He (she) did something his (her) mother liked very much and she decided to do something nice.

Which would be best for Mike (Mary) or which would he (she) like better?

Would it be better or would Mike (Mary) like it better

CHILD DEVELOPMENT

<i>Items</i>	<i>Category</i>
If his (her) mother said, "Santa Claus will be good to you."	Supernatural Force
If his (her) mother said, "I'll take you on a trip to visit someone you like."	Visiting
If his (her) mother gave him (her) a toy (or something nice).	Material Reward
If his (her) mother hugged him (her).	Physical Affection
If his (her) mother said, "You have been a very good boy (girl)."	Praise
If his (her) mother said, "You have made me very happy."	Emotional Appeal
If his (her) mother said, "We'll all get together and do something."	Centrality
If his (her) mother said, "I love you very much."	Verbal Love

RESULTS

The reliability of the children's responses can be approached in two ways: First, how well did they agree with themselves from session to session? Second, within a particular session, how consistent were the children from the viewpoint of logic?

The extent of agreement from session to session on each of the four scales is given in Table 1. It can be seen that the reliability of the responses is somewhat in the predicted direction. The highest extent of agreement is for the O scale, the task of which was ordering objects along a physical dimension. It should be noted that, where the scales involved judgments of a more social nature, the extent of agreement decreased.

The mean percentage of agreement is also given, mainly for illustrative purposes. Why the reward scale should indicate much less agreement than the punishment scale is not known. It is interesting to speculate, however, that children may have more experience with punishments than rewards, leading to more stable judgments. It should be noted here that the percentage of agreement for the P scale (.48) agrees favorably with the identical scale used by Clifford and Wischner (1).

TABLE 1
EXTENT OF AGREEMENT FROM SESSION I TO SESSION II ON EACH
OF THE FOUR SCALES

	<i>Scale O</i>	<i>Scale F</i>	<i>Scale P</i>	<i>Scale R</i>
Mean number of agreements	17.78	15.56	13.61	8.83
Standard deviation	5.34	4.45	3.59	3.10
Mean percentage of agreement63	.55	.48	.31

EDWARD CLIFFORD

TABLE 2

INTERSCALE SIGNIFICANCE OF THE DIFFERENCES BETWEEN MEAN NUMBER OF AGREEMENTS FROM SESSION I TO SESSION II (*t* Scores)

	<i>Scale F</i>	<i>Scale P</i>	<i>Scale R</i>
Scale O	1.313	2.673**	7.160***
Scale F		1.402	5.098***
Scale P			4.156***

** Significant at .01 level.

*** Significant at .001 level.

The significance of the interscale differences between the means for the four scales was assessed using *t*. These data are given in Table 2.

The extent of agreement for the R scale is significantly poorer ($p = .001$) than for the other three scales. The O scale demonstrates significantly higher agreement from session to session than the P scale. On the basis of these results, it would seem clear that the extent of agreement can be significantly increased with the introduction of a scale that can be physically ordered. This finding should not be too surprising, since a physically ordered scale is the historical basis for the paired comparisons method.

Another approach to reliability is to examine how consistent the children were in making the judgments. To do this, each record was examined for circular triads. (A circular triad is an inconsistency in judgment of the order A is greater than B, B is greater than C, C is greater than A.) Following Edwards (2), for each scale there are 20 possible circular triads per child, or, if all the children were inconsistent on a particular scale, a maximum of 360 circular triads.

TABLE 3

THE NUMBER OF CIRCULAR TRIADS IN EACH OF THE FOUR SCALES IN SESSION I AND SESSION II

<i>Scale</i>	<i>Session</i>	<i>Total Triads</i>	<i>Mean Triads</i>	<i>SD</i>
O	I	89	4.94	5.32
O	II	123	6.83	5.97
F	I	199	11.06	5.24
F	II	119	6.61	6.02
P	I	156	8.67	4.50
P	II	144	8.00	5.55
R	I	125	6.94	4.12
R	II	148	8.22	3.71

CHILD DEVELOPMENT

Table 3 presents the number of circular triads in each of the scales in both sessions. It can be seen that even in the "most reliable" scale (as measured by extent of agreement from session to session), a large number of circular triads are present.

When the scales are compared and the differences between the means examined, none of the mean differences between scales is significant for the second session. In the first session mean differences between Scale O and Scales F ($p = .001$) and P ($p = .05$) and the difference between Scale F and Scale R ($p = .01$) are significant. For the first session, then, the children had fewer circular triads on the physically ordered scale than on the other scales administered; while there is a tendency for this to continue in the second session, the differences obtained are not significant.

While the range in age is admittedly narrow in the present sample, it is of some interest to examine the possible relationship between the number of circular triads and age. Table 4 presents this information as well as a comparison of the number of circular triads across scales in terms of rank order correlation coefficients.

TABLE 4
RANK ORDER CORRELATIONS (BASED ON THE NUMBER OF CIRCULAR TRIADS) BETWEEN SCALES AND AGE FOR SESSION I AND SESSION II

	Age	Scale O	Scale F	Scale P
SESSION I				
Scale O	-.54*			
Scale F	-.10	.51*		
Scale P	-.20	.35	.54*	
Scale R	-.15	.04	.22	.27
SESSION II				
Scale O	-.21			
Scale F04	.10		
Scale P	-.31	.19	-.16	
Scale R50*	.02	-.05	-.41

* Significant at .05 level.

In the first session there is a tendency for the number of circular triads to be associated with the age of the child. The negative correlations obtained indicate the tendency of the older child to have fewer circular triads. Only one of the correlations obtained is significant, that for the O scale, but this fails to achieve significance in the second session. There also seems to be a minimal relationship between the number of triads on any one scale and the number of triads on another. These data, taken in conjunction with the data concerning the extent of agreement with oneself from session to session, tend to indicate, therefore, that the child at this age level may not

EDWARD CLIFFORD

TABLE 5

THE NUMBER OF TIMES EACH ITEM WAS CHOSEN OVER EVERY OTHER
ITEM FOR BOTH SESSIONS

SESSION I		SESSION II	
Items	Freq.	Items	Freq.
S C A L E O			
House	101	Car	92
Car	96	House	88
Table	84	Table	82
Chair	79	Chair	58
Pot	49	Pot	57
Dish	40	Pin	45
Spoon	35	Dish	44
Pin	20	Spoon	38
$Rho = .85$			
S C A L E F			
Grapes	73	Grapes	76
Cherries	70	Bananas	66
Apples	69	Apples	64
Bananas	63	Grapefruit	61
Pears	59	Peaches	61
Peaches	59	Pears	60
Oranges	58	Oranges	58
Grapefruit	53	Cherries	58
$Rho = .43$			
S C A L E P			
Deprivation	74	Spanking	75
Spanking	74	Supernatural Force	73
Supernatural Force	69	Deprivation	69
Desertion	61	Shaming	65
Scolding	61	Scolding	62
Shaming	55	Isolation	58
Isolation	55	Desertion	53
Withdrawal of Love	55	Withdrawal of Love	49
$Rho = .75$			
S C A L E R			
Supernatural Force	74	Material Reward	75
Praise	70	Supernatural Force	72
Material Reward	69	Visiting	69
Centrality	62	Physical Affection	64
Visiting	61	Verbal Love	60
Physical Affection	59	Praise	59
Emotional Appeal	55	Emotional Appeal	55
Verbal Love	54	Centrality	50
$Rho = .36$			

CHILD DEVELOPMENT

be capable of making the necessary judgments regardless of the type of scale used, although there is some indication that a scale based on some preordered basis would be more effective.

For each of the scales the number of times each item was chosen over every other item was determined and the items were assigned a rank order. These data are given in Table 5.

Despite relatively low individual reliabilities, Scales O and P show rather surprisingly high correlations. The rho for Scale P compares favorably with that obtained by Clifford and Wischner (1). Why these two scales should demonstrate higher relationships from session to session is unknown. One may speculate, however, that the items that compose these scales may very well be more within the range of the child's experience than the items involved in either the F or R scales.

CONCLUSIONS

Eighteen preschool children were given four scales verbally in a paired comparison procedure. The extent of agreement from session to session was significantly higher for the scale containing items on a physical dimension than for the scales containing punishment or reward items. Consistency within a session for any of the scales was not extensive, casting doubt on the use of this procedure with this age group.

This study indicates the need for caution in utilizing the paired comparisons procedure with preschool children. A need is also demonstrated for further work along an age dimension to determine at what level this technique can be used with adequate reliability. In addition, a variety of scales need to be examined in order to determine what types of scales can best be adapted to this procedure.

REFERENCES

1. CLIFFORD, E., & WISCHNER, G. J. The relative severity of different kinds of parental punishments as evaluated by preschool children. Paper read at Midwest. Psychol. Ass., Chicago, April, 1951.
2. EDWARDS, A. I. *Techniques of attitude scale construction*. New York: Appleton-Century-Crofts, 1957.
3. GUYETTE, R. Relationships between children's, parent's and teacher's perceptions of the severity of punishments. Unpublished master's thesis, Univer. of Massachusetts, 1957.
4. KOCH, HELEN L. Popularity in preschool children: Some related factors and a technique for its measurement. *Child Developm.*, 1933, 4, 164-175.
5. LIPPITT, ROSEMARY. Popularity among preschool children. *Child Developm.*, 1941, 12, 305-332.

A DEVELOPMENTAL ANALYSIS OF SOME BODY IMAGE AND BODY REACTIVITY DIMENSIONS¹

SEYMOUR FISHER and RHODA LEE FISHER*

Baylor University College of Medicine

It was demonstrated in previous work (8, 9, 10) that an individual's attitudes toward his body (body image) provide a basis for understanding important aspects of his behavior. Body image measures were able to predict such diverse phenomena as level of aspiration, site of psychosomatic symptomatology, ability to tolerate stress, and behavior in small groups (10). There was particular success in utilizing body image concepts to predict characteristic individual patterns of physiological reactivity. In an initial series of studies (10) it was shown that the degree of definiteness ascribed by the individual to his body boundary was significantly linked with the ratio of body exterior (skin and muscle) to body interior (heart and viscera) reactivity he manifested. A definite and firm concept of one's body boundary is accompanied by a relatively higher degree of exterior than interior response. An indefinite concept is correlated with just the opposite pattern of reactivity. The over-all findings were such as to suggest that the individual's attitudes toward various sectors of his body could influence the reactivity of those sectors. In order to test the generality of this formulation beyond the interior-exterior differentiation a number of experiments (5, 6, 7) were undertaken which concerned the influence of attitudes toward the right and left sides of the body upon the relative

* Baylor University College of Medicine, Houston, Texas.

¹ This study was supported mainly by USPH Career Investigator Grant M-1578.

This research was also supported in part by the U.S. Air Force through the Air Force Office of Scientific Research of the Air Research and Development Command under contract number AF 18 (603)-79.

CHILD DEVELOPMENT

reactivity of the two sides. Various body image measures were developed for determining the relative size or strength attributed by a person to the right vs. left sides of his body. Further, a GSR technique was devised for measuring the actual reactivity of the two body sides. This GSR technique involved taking balanced recordings from a series of homologous left-right sites and comparing the relative response amplitudes from each pair of sites. Unequivocal results were obtained which indicated that in right-handed² adults there is an optimum norm gradient such that the left side is more reactive than the right side. This gradient significantly typifies Ss who in their body image evaluations clearly distinguish the right side of the body as larger or stronger than the left side. Persons who do not make such a clear body image distinction in this respect either fail to manifest a GSR right-left gradient or are more reactive on the right side. It was also found in two separate studies (6, 7) that a left directional GSR gradient occurred significantly more often in individuals who had a generally well integrated and stable body image. Those with poorly integrated body images showed either right directional reactivity or no gradient at all. It was concluded that a mature body image was a necessary foundation for the establishment of a left directional gradient.

The finding of a link between body image indices and a clear-cut dimension of body reactivity quickly suggested the possibility that it might be utilized as a simple model for studying body image-body reactivity relationships developmentally. An unusually good opportunity seemed to present itself for determining whether there are meaningful sequences of relationship between body image and a dimension of body reactivity when viewed over a broad time perspective. It was considered that such data would help to clarify not only origins of body image attitudes but also the manner in which they become linked with patterns of body reactivity. Accordingly, the formal objective of the present study was to secure various body image indices and also the right-left GSR measure from a range of Ss differing in chronological age and to observe whether there were definable developmental variations and trends.

PROCEDURE

Body Image Measures

The following body image measures were obtained from each S:

1. *Aniseikonic lens comparison of left and right body sides.* This measure has been described in detail elsewhere (5). The S is asked to don a pair of aniseikonic lenses (4.00 per cent power) and told that the lenses will produce distortions in his visual field. When it is apparent that he is

² In left-handed Ss there is an absence of relationship between GSR right-left gradients and body image indices of attitudes toward the right vs. left body sides. A detailed consideration of this point may be found elsewhere (5).

actually experiencing dramatic distortions, he is asked to make a series of comparisons of the sizes of the fingers of his two hands. It is emphasized that the differences involved are very small and that even the most minute size discrepancy between any two fingers should be reported. The objective is to communicate to the *S* that his judgment is not reality bound and that he should try to experience a difference whenever possible. It is assumed that in such a judgmental setting the individual with a clear body image differentiation between the right and left body sides would be inclined to perceive differences in the finger sizes and that the differences would be consistently in a specific direction. A *S*'s score for his aniseikonic judgments may vary from the extreme of seeing no differences between any of the sets of fingers to the other extreme of perceiving each of the five fingers on a given body side as larger than the corresponding fingers on the opposite side.

2. *Puppet representations of the right and left body sides.* This is a new measure intended to tap differential attitudes toward the left and right body sides in terms of symbolic acting out. Each *S* was asked to place identical cylindrical lengths of paper over the middle finger of each hand and to imagine that they represented puppets. He was then to give a name to each and to make up a brief story about the two of them. A total of four such stories, each involving two new puppet characters, was requested from him. The assumption underlying this procedure was that *S*s would identify each puppet with a given body side and therefore in describing them would symbolically express the qualities they associate with each side.

The right vs. left variable that was selected out for study by means of the puppet technique had to do with masculinity-femininity. It concerned the relative degree of masculinity vs. femininity attributed to each body side. This index was computed simply by counting the number of masculine vs. feminine names assigned to each puppet on the four different occasions that new names were offered by the *S*. For any given side the maximum number of points obtainable for either of the sex identities was 4 and the minimum was 0.

3. *Measure of body image boundary definiteness.* This index which is referred to as the Barrier score is computed from responses to the Rorschach blots. It has been demonstrated (10) that the Barrier score represents an evaluation of the degree to which the individual pictures the boundaries of his body as definite and firm as contrasted to indefinite and weak. The scoring of any individual's Rorschach protocol for this variable is based on a count of the number of percepts in which unusual covering, containing, decorative, or protective qualities are ascribed to the bounding peripheries. Such responses as the following would exemplify the category: "cave with rocky walls," "knight in armor," "animal with striped fur," "fort with cement walls." Detailed scoring norms may be found elsewhere (10). The larger the Barrier score the greater is the definiteness of the body image boundaries.

CHILD DEVELOPMENT

This measure was included in the present study because it has proven itself to represent a very fundamental body image dimension. It provides a possibility of relating developmental variations in GSR gradient not only to specific attitudes toward the right and left body sides but also to a more generalized body image variable.

4. *Thematic Apperception Test.* Cards 1, 2, 5, 6BM, and 7BM of the TAT were administered for exploratory purposes and also to obtain data concerning the types of roles assigned to the parental figures. It was felt that data of this sort might provide historical perspective for viewing the results obtained from the body image measures. Most of the pictures included in this TAT series contain representations of the parental figures and provide information about modes of relationship with them. It was the intent to focus on how body image-body reactivity right-left patterns might have evolved from having learned certain roles and values in interaction with the parental figures.

GSR Measure of Right-Left Reactivity

This technique for determining the character of an individual's left-right GSR gradient has already been presented in detail elsewhere (5). Therefore, it will be only briefly summarized here. The GSR method was of the exosomatic type. Recordings from the right side and left side were made simultaneously on separate channels against a relatively inactive upper arm. Such recordings were taken from four different homologous sites. One could then compare amplitudes of responses at homologous points at a given time. The criterion of right vs. left directionality of response was based upon comparison of these amplitudes. If the differences occurring for a specific site were such that there were twice as many favoring one side as the other and if at least eight such differences were present, then it was judged that a directionality for that site had been demonstrated. In addition, when it was found that the directionality was in favor of one side for three out of the four sites, it was considered that there was a significant GSR gradient in that direction. When all four sites showed a majority of the differences in the same direction, a less stringent criterion of significant directionality was applied in evaluating any single site. It was then necessary for only five of the differences at any one site to be in the given direction in order to designate that site as showing definite directionality.

Subjects

There were 119 Ss of which 53 were boys and 66 girls. They were all right-handed. Their ages ranged from 3 to 17 years. They were paid a fee for participating. The younger Ss were recruited by contacting families and neighborhood mothers' groups known to the experimenters. The older Ss were recruited in the main from Boy Scout and Girl Scout groups. As defined by the occupational level of the father, the Ss were almost without exception of middle class backgrounds.

SEYMOUR FISHER and RHODA LEE FISHER

TABLE I

NUMBERS AND PERCENTAGES OF SUBJECTS MANIFESTING VARIOUS
GSR GRADIENT PATTERNS AT DIFFERENT AGE LEVELS

Age	Right Reactive		Left Reactive		No Gradient	
	N	%	N	%	N	%
3-6*	10	67	0	0	5	33
7-8	7	44	1	6	8	50
9-10	7	29	5	21	12	50
11-12	2	25	3	37.5	3	37.5
13-14	9	27	13	40	11	33
15-17	4	17	10	44	9	39

* Adjacent age levels have been combined in order to provide an *N* at most levels which is at least moderately substantial.

RESULTS

GSR Gradient Variations

When the directionalities of the GSR gradients occurring at the various age levels are reviewed, one finds some striking trends. Table I presents these trends. It can be seen that among the younger *Ss* there is either an absence of a gradient or greatest reactivity occurs on the side of the dominant hand. Thus, in the age range 3 to 6 there are no instances of left reactivity, but 67 per cent of the group are right reactive and 33 per cent have no gradient. This age level stands in great contrast to the 15 to 17 group in which 44 per cent are left reactive, 17 per cent right reactive, and 39 per cent without a gradient. There is a very slow increase in the incidence of left reactivity in the sequence of ages from 3 to 10. It first begins to become a noteworthy phenomenon in the age range 11 to 12. Then, in the ages that follow it seems to stabilize so as to constitute about 40 per cent of the total reactivity patterns. The fact that in a previous study (5) involving adults (median age 25) 48 per cent of the group was left reactive suggests that an incidence somewhere in the 40 per cent range is rather constant beyond the age of 12.

The incidence of right reactivity declines very sharply over the age range from 3 to 9. It then seems to level off at an incidence level in the 20 per cent range from ages 9 to 14 and falls further to 17 per cent in the 15 to 17 group. It is interesting that in an adult population (5) it was also found that 17 per cent of the group were right reactive.

One does not detect any clear trends concerning the incidence of "no gradient" patterns in the various age groups. There is perhaps a tendency for *Ss* in the younger age groups to be characterized less frequently than the older *Ss* by a definite gradient, but such tendencies do not even approach statistical significance.

CHILD DEVELOPMENT

In scanning the data it seemed logical to split the over-all age sequence into a preadolescent period and an adolescent period. That is, it appeared that up to adolescence the GSR gradient was mainly associated with hand dominance, and it was only with the onset of adolescence that a new dimension (left reactivity) entered into the picture. It was therefore decided to dichotomize the Ss into a preadolescent grouping consisting of those 3 to 12 years of age and an adolescent group embracing all Ss 13 to 17. The relatively late age of 13 was arbitrarily set as the lower age for the adolescent group in order to make certain that the majority of Ss in this category would have actually encountered the traditional problems of adolescent adjustment. When the incidence of left reactivity vs. other patterns of reactivity is compared in the preadolescent and adolescent groups by means of chi square, it is found that left reactivity is highly significantly (.001 level) greater in the adolescent group. One also finds that when the incidence of right reactivity vs. other patterns is compared in the two groups that right reactivity is significantly (.05 to .02 level) greater in the preadolescent group.

It should be parenthetically noted that there are no sex differences in the incidence of the GSR patterns in any of the groups.

Aniseikonic Lens Judgments

The next phase of the analysis of the data concerned the link between the aniseikonic lens measure and GSR patterns in the preadolescent group and in the adolescent group. It may be first noted that the median number of fingers of the right hand judged larger than those of the left hand was 3 in each of the following age grouping: 5 to 8, 9 to 12, 13 to 16. This median was also found in an adult group previously studied (5). There were no sex differences in aniseikonic judgments in any of the age groups.

A chi square analysis of the relationship of the aniseikonic judgments to GSR reactivity patterns in the preadolescent group gave entirely chance results. There was no apparent link between the degree to which the two sides of the body were clearly differentiated and the incidence of any given type of GSR pattern. However, in the adolescent group a chi square test did reveal significantly greater (.05 to .02 level) incidence of left GSR reactivity among those above the median in aniseikonic differentiation of the right and left body sides than among those below the median in this respect. This finding in the adolescent group cross-validates similar results previously observed in a group of adult Ss (5). It would appear, then, that in the period previous to adolescence the GSR right-left pattern is not linked with the aniseikonic body image index, whereas during adolescence it does become significantly related.

Puppets

The relationships of the puppet scores to GSR directionality at the 13 to 17 level were examined. There were previous findings (6) which suggested that the left directional gradient might be most likely to occur

in individuals who clearly differentiated the sexual characteristics of the two sides in such a fashion as to view the left side as more feminine than the right side. Chi square indicated that there were only chance relationships between GSR directionality and the ratio of right side masculinity to left side masculinity or the ratio of right side femininity to left side femininity. However, when Ss were categorized into those above and below the total group median with regard to number of masculine names assigned to the left hand alone, it was found by chi square that those below the median in this respect were significantly (.01 to .001 level) more often left GSR reactive than those above the median. The same significant (.01 level) results were noted when the boy's group was analysed separately. While the results were in the predicted direction in the girl's group, they were not significant.

Another analysis was undertaken which involved the entire group and concerned Ss above and below the median number of feminine names assigned to the left hand alone. A chi square of 3.9 (.05 level) indicated that Ss above the median in left femininity were more likely to be left GSR reactive than Ss below the median in this respect. There was an absence of significant relationship between GSR gradient and the masculinity or femininity ascribed to the right side alone.

Boundary Definiteness

The Barrier score, which is derived from the Rorschach protocols, revealed no gross sex difference in the age range 5 to 12. It did demonstrate a significant tendency, within the 13 to 17 group, for the girls to have more definite body image boundaries than the boys. The chi square value fell at the .01 level.

The mean Barrier scores of the 5 to 12 Ss in the various GSR directionality categories were characterized by only chance differences. However, within the 13 to 17 group the mean Barrier score of the left reactive Ss was significantly higher (.05 level) than the combined Ss who were either right reactive or manifested no gradient at all.

Thematic Apperception Test

Analysis of the TAT stories was devoted primarily to determining what sorts of roles were ascribed to the parental figures. Because a previous study (6) indicated that GSR directionality was related to the clearness with which the individual had formulated the male and female roles, it seemed logical to evaluate the roles he assigned to parental figures that were pertinent to sexual differentiation. A separate analysis was undertaken for each sex.

The character of the five TAT cards (1, 2, 5, 6BM, 7BM) which were utilized dictated to a considerable degree the sorts of measures of ascribed parental role that could be derived. Three measures were constructed for the girls:

CHILD DEVELOPMENT

1. *Concept of father on 7BM.* A simple evaluation was made as to whether the father figure in 7BM was or was not given a status higher than that of the younger figure in the picture. That is, was father perceived as having a traditionally forceful masculine role? Higher status was defined in terms of being designated as a senior person (e.g., professor vs. student) or as one who is controlling or guiding or protecting the younger figure.

2. *Concept of mother on 6BM.* The same evaluation criterion was used as that applied to judging father's role on 7BM.

3. *Use of sexual role imagery in a parental context.* This measure was taken from the response to card 2 which involves three figures who are frequently seen as father, mother, and daughter. The measure was based on the presumption that, if one's parental figures are experienced as people with well defined and comfortable sexual roles, they constitute a model of permission for oneself to have sexual goals. Note was taken therefore of whether the younger figure in the picture was or was not described as having a heterosexual love relationship with some other person.

The measure of father status derived from picture 7BM proved to have a significant relationship to GSR directionality. Chi square indicated that girls who depicted the father as having a clear-cut dominant role were more likely (.01 to .001 level) to be left reactive than those not portraying him in such a fashion.

The measure of mother figure status taken from picture 6BM proved not to be operationally feasible. It was eliminated because with few exceptions most of the stories to be judged provided little information concerning the status relations of the mother figure and the younger figure.

The sexual role imagery index based on picture 2 demonstrated itself to be meaningfully tied in with GSR directionality. A chi square test which was significant at the .05 level indicates that girls who attributed heterosexual aims to the young figure in the picture were more often left reactive than those who could not see her with such aims.

The same TAT measures were applied to the boy's group. However, only chance relations were found between them and the GSR directionality categories. This may simply reflect the fact that there were only eight left reactive individuals in the boy's 13 to 17 group which is a rather limited sample against which to compare the 19 Ss in the other GSR category.

Relationships Among Body Image Measures

The body image measures significantly related to GSR directionality were compared with each other. There were three such measures: the aniseikonic judgments, the puppet masculinity-femininity index, and the Barrier score. All comparisons involved the 13 to 17 age range. Chi square showed the aniseikonic judgments to be related to degree of puppet left-hand femininity (at the .05 to .01 level) in the expected direction. That is, those Ss most clearly differentiating the right hand as larger than the left

SEYMOUR FISHER and RHODA LEE FISHER

also assigned a relatively high degree of femininity to the left hand. A borderline relationship (.10 to .05 level) was found between anisikonic judgments and the Barrier score. But the tendency was for those with higher Barrier scores to differentiate most clearly between the right and left sides. Only a chance relation was found between the puppet measure of left-hand femininity and the Barrier score.

INTERPRETATIONS

The data demonstrate a meaningful developmental pattern with regard to right-left reactivity gradients. It would appear that previous to adolescence greatest reactivity occurs either on the side of the dominant hand or else there is no consistent difference between the two body sides. In this preadolescent range one finds that none of the body image measures are linked with GSR reactivity (i.e., right reactivity vs. no definite gradient). But in the adolescent period GSR asymmetry is significantly linked with a number of body image indices. Thus, it is related to the clearness with which the size attributes of the right and left sides can be distinguished, the degree of masculinity-femininity assigned to the left side of the body, and the definiteness attributed to the over-all body image boundaries.

What is the meaning of this shift that occurs at adolescence? One should consider the possibility that changes in body structure and hormonal balance associated with puberty play a significant part in the right-left reactivity shift and even influence the body image attitudes that are correlated with right-left reactivity patterns. Since measures pertaining to body structure and hormonal balance were not included in the present study, it is impossible to rule out such a contingency. However, the explanatory model which will be offered here concerning the shift assigns most importance to socialization and learning variables. This speculative model starts out with the assumption that previous to adolescence the reactivity of the two sides is governed primarily by factors associated with lateral dominance. It has, of course, been shown (11) that there are differences in the motor reactivity of the two body sides which are correlated with lateral dominance. The possibility presents itself that definite right GSR directionality reflects the fact of well established right lateral dominance which results in a greater ratio of excitation being channeled to the right than to the left side. That is, greater responsiveness would occur on the right side which is most prepared or set for reaction. The absence of a GSR gradient might mirror, then, a relatively low degree of established lateral dominance and a decrease in the differential between the two body sides in their degree of preparatory set for response. It may be hypothesized that with the coming of adolescence right vs. left GSR reactivity would no longer be governed by such a simple principle. There might be a reorganization that would introduce new complexity into the process. The left-right response gradient would begin to be influenced by the symbolic values assigned to each side.

CHILD DEVELOPMENT

Data obtained by means of the puppet technique suggest that especially prominent among the symbolic differences which might take on increased import at adolescence are those pertaining to masculinity-femininity. It was shown that at the time of adolescence the greater the number of feminine names assigned to the left-hand puppet the greater the probability that the GSR gradient will be left directional. Perhaps there develops during this period, which is so characterized by attempts at sex role structuring (2, 3), a sharper body image distinction between the two body sides relative to the masculine or feminine values attributed to each. The possible significance of sex role factors in the emergence of new patterns of GSR right-left reactivity is underscored by the results obtained from the TAT measures. It will be recalled that in the female adolescent group left GSR directionality was significantly more frequent among those who pictured the father figure in a traditionally strong male role than those who did not do so. Also, it was noted that in this group left directionality occurred most often among those who could, in TAT terms, identify with an active heterosexual role. In a related vein, it was found in a previous study (6) that left GSR directionality was significantly more frequent among individuals drawing the male figure larger than the female figure than among those not making such a distinction.

The question arises as to what mechanisms might be involved in the relationship of sex role differentiation to GSR gradient which seems to evolve at the time of adolescence. It was originally proposed (5) that in adults the GSR gradient represented the fact that one side of the body had been assigned major responsibility for being set or ready to deal with important stimuli. It was suggested that in order for this set to be functionally effective the body side involved had to be maintained at a level of inhibition which would prevent its response to transitory surges of excitation. That is, it had to be prepared for quick response in the face of important matters but be relatively unresponsive to unimportant events that might continually interfere with its long term alertness functions. The hypothesis was proposed that in order for this set to be maintained there was a tendency to shield that side from excitation and to channel the excess to the other body side. Thus, in a right-handed individual this would mean that the right body side would be held in a state of readiness and that this set would be protected from disrupting fluctuations of excitation by providing for their diversion to the left side. The findings of the present study imply that up to adolescence there is no well worked out mechanism of this sort. Rather, in the preadolescent period the preparatory set of the right side seems to make it a channel into which all excess excitation is turned. It may be assumed that more complicated means of distributing excitation between the two body sides arises at adolescence and that this more complex system is at least in part founded upon the individual having proceeded from a relatively simple body image concept which does not distinguish the two body sides in terms of sexual values to a more complex

concept in which there is such a distinction. It is not clear at this point in what specific fashion the more complex body image differentiation would aid in establishing a more complex excitation allocation system. Perhaps, the mere fact of assigning differentiated significances to the two sides facilitates their functioning in a specialized manner at other levels. The data do not permit more than speculation regarding such points at present. A variety of matters remain unexplained. Primary among these is the question why it is the number of feminine names assigned to the left-hand puppet which proved best able to distinguish the GSR reactivity groups. Why is it that it is left-hand designations rather than the right-hand designations which are significant? Considerable more work will be required to elucidate this problem.

Although great emphasis has been placed on the importance of sex role integration in the development of the right-left reaction gradient, one must not overlook the fact that a significant relationship was also found between the Barrier score and gradient characteristics. It has been shown (10) that the Barrier score taps a level of body image organization which occurs far earlier than adolescence and involves the very fundamental problem of establishing boundaries to demarcate one's body off from the rest of the world. The present findings indicate that adolescents who have definite body image boundaries are more likely to be left reactive than adolescents who manifest indefinite boundaries. Therefore, the question arises whether the development of the left directional gradient is not also influenced by how well the individual works out other body image problems which are probably prior to that of the organization of sexual attributes.

As one surveys the total pattern of the data, it is apparent that body image attitudes and associated body reactivity patterns fall into meaningful sequences over a developmental range. One can see that the relationships between body image and body reactivity in adulthood are the outcome of an orderly series of preceding events. The evidence is now definite that body image factors play a significant part in an area of behavior which has usually been isolated under the separate rubric of "physiological response." The application of a body image frame of reference has made it possible to demonstrate with unusual clarity that there are long term links between body reactivity patterns and attitudes which are probably learned from social interaction (especially with the parents). It should also be pointed out that the application of a body image frame of reference has made it possible to predict the existence of new dimensions of body reactivity which had not been previously defined. It has enabled the prediction of the right-left pattern and also a body interior vs. body exterior reactivity pattern which is discussed elsewhere (10). The isolation of two such gradients opens interesting possibilities. It now seems likely that there are still other gradients and that they also would be found to be related to body image variables. The fact that the first two gradients studied have involved obvious spatial dimensions of the body suggests that it would be fruitful

CHILD DEVELOPMENT

to investigate other analogous dimensions. For example, there may be a meaningful dimension relating to the reactivity of the front of the body vs. the back of the body or one involving the head vs. the body. There are certainly many clinical observations (4) which indicate that these spatial dimensions are quite meaningful in a body image sense.

The present study has provided still further elaboration to a series of constructs concerning body image which the senior author is attempting to formulate. It is the ultimate intent to build up a cohesive theory which will isolate the main dimensions of body image and define the influence they have upon various levels of behavior.

SUMMARY

This study concerns itself primarily with a developmental analysis of the relationship between body image concepts of the right and left body sides and the relative GSR reactivity of the two sides. It is an extension of previous work which sought to demonstrate that the physiological reactivity of given body sectors is linked with body image attitudes toward those sectors. A total of 119 right-handed Ss in the age range 5 to 17 years were tested with the following battery of measures:

1. A method utilizing aniseikonic lenses to obtain projective judgments of the relative sizes of the two body sides.
2. An evaluation by means of right-hand and left-hand puppets of the symbolic values assigned to the two body sides.
3. An index of degree of body image boundary definiteness derived from Rorschach responses.
4. Thematic Apperception Test indicators of attitudes toward the parental figures.

In addition, simultaneous balanced GSR recordings were taken from the two body sides in such a way as to determine whether the individual was characteristically left reactive, right reactive, or without a definite gradient.

The principal findings may be reduced to the following:

1. In right-handed Ss left directional reactivity is rare until adolescence. Up to adolescence either right reactivity or absence of a gradient is most characteristic. At the time of adolescence the incidence of left reactivity rises to approximately 40 per cent.
2. There are only chance relationships previous to adolescence between body image measures and right-left reactivity patterns, but during adolescence they become significantly linked. A left directional gradient then becomes typical of the S whose body image is marked by clear right-left differentiation and definite boundaries. Right directional reactivity or absence of a gradient is found most frequently in the individual with a poorly differentiated and poorly bounded body image.

3. There are suggestions that the development of effective differentiation in the right-left body image axis is to some degree a function of having parental models who exemplify well defined sex roles.

4. A more complex system for the disposition of excitation seems to arise at adolescence. It has been hypothesized that a prime prerequisite of this system is that the individual learn to make a clear distinction between the male and female roles and on this basis to differentiate the two body sides within a context of masculine and feminine values.

The general findings indicate, then, that the body image-body reactivity relationships found in adults are the final outcome of a preceding developmental process with predictable features.

REFERENCES

1. BAUER, R. W., & WEPMAN, J. M. Lateralization of cerebral functions. *J. Speech Hearing Disorders*, 1955, 20, 171-177.
2. DENNIS, W. The adolescent. In L. Carmichael (Ed.), *Manual of child psychology*. New York: Wiley, 1946. Pp. 633-666.
3. ERIKSON, E. H. *Childhood and society*. New York: Norton, 1950.
4. FENICHEL, O. *The psychoanalytic theory of neurosis*. New York: Norton, 1945.
5. FISHER, S. Body image and asymmetry of body reactivity. *J. abnorm. soc. Psychol.*, in press.
6. FISHER, S. Body reactivity gradients and figure drawing variables. *J. consult. Psychol.*, in press.
7. FISHER, S., & ABERCROMBIE, J. The relationship of body image distortions to body reactivity gradients. *J. Pers.*, in press.
8. FISHER, S., & CLEVELAND, S. E. The role of body image in psychosomatic symptom choice. *Psychol. Monogr.*, 1955, 69, No. 17. (Whole No. 402)
9. FISHER, S., & CLEVELAND, S. E. An approach to physiological reactivity in terms of a body-image schema. *Psychol. Rev.*, 1957, 64, 26-37.
10. FISHER, S., & CLEVELAND, S. E. *Body image and personality*. Princeton, New Jersey: Van Nostrand, 1958.
11. JASPER, H. H., & RANEY, E. T. The physiology of lateral cerebral dominance. *Psychol. Bull.*, 1937, 34, 155-165.
12. MEAD, M. *Coming of age in Samoa*. New York: Morrow, 1928.

CHILD DEVELOPMENT

13. MEAD, M. *Growing up in New Guinea*. New York: Morrow, 1930.
14. THASS-THIENEMANN, T. Left-handed writing. *Psychoanal. Rev.*, 1955, 42, 239-261.
15. WITKIN, H. A., LEWIS, H. B., HERTZMAN, M., MACHOVER, K., MEISSNER, P. B., & WAPNER, S. *Personality through perception*. New York: Harper, 1954.
16. WOLFF, W. *The expression of personality*. New York: Harper, 1943.

YOUNG CHILDREN'S DISCRIMINATIONS OF PARENT AND CHILD ROLES^{1, 2}

WALTER EMMERICH*

University of Colorado

There is increasing interest in viewing interpersonal relationships as social systems composed of role elements. In a theoretical analysis of the socialization process, Parsons (10) applies this approach to changes in parent-child relationships and the child's personality structure at different periods of development. In broadest outline, his view is that: (a) each phase of socialization is initiated by an increased differentiation of the social structure into which the child is being socialized; (b) under the impact of each "new" social structure, there are certain changes in the motivations, cognitions, and performances of the child; (c) there is an ordered sequence of social structures during socialization and a corresponding developmental sequence of needs and role conceptions in the child; (d) each reorganization of needs and roles in this sequence arises from a discrimination (bifurcation) on a specific dimension to which the child was previously insensitive; (e) the dimensions used to discriminate needs and roles are derived from the Pattern Variables (11), postulated to be a basic and exhaustive set of binary choices made in any action sequence by the mature adult.

This study investigates several aspects of Parsons' theory concerning the development of parent and child role concepts in the young child. The mother, father, daughter, and son are treated as four basic role types in the family social structure. It is hypothesized that the discrimination of

* Department of Psychology, University of Colorado, Boulder.

¹ This study was supported by a grant from the Council on Research and Creative Work of the University of Colorado.

² The author wishes to express his gratitude to the following persons, whose cooperation made this study possible: Assistant Superintendent Melvin Wiesley, Principal Keith Dillman, and the teachers of the Martin Park Elementary School, Fairview District, Boulder, Colorado; Mrs. Jean Goodnow, Mrs. Mary Jane Eff, and the parents of the Presbyterian Nursery School, also of Boulder, Colorado.

CHILD DEVELOPMENT

parent and child roles is made on the basis of relative power, and that the discrimination of female and male roles is made on the basis of their relative emphases upon expressive and instrumental functions, respectively. Thus, it is expected that each basic role type will be described in terms of the four combinations of the power and function dimensions:³ mother—high power, expressive; father—high power, instrumental; daughter—low power, expressive; son—low power, instrumental. However, these categorizations represent family roles as viewed by the adult, being an end-point in a hypothesized developmental sequence of categorizations by the child. Discrimination of parent and child roles is said to occur first, beginning during the first year and increasing during the second and third years.⁴ Discrimination of sex roles on the function dimension is hypothesized to occur next, beginning in about the third year and increasing throughout the preschool period.

Parsons provides only implicit definitions of role categorizations based upon the power and function dimensions (10). It is therefore necessary to define these dimensions explicitly before deriving a set of behaviors representative of each of the four role types. Relative power shall be defined in terms of how much control each person has over the outcome of an interaction sequence. For example, A's praise of B implies that A is allocated the power either to grant or withhold what B wants; and reciprocally, B's seeking praise from A implies that whether or not B gets what he wants is contingent upon A's action. The instrumental-expressive dimension is more difficult to define. Generally, the instrumental function is task-oriented ("adaptive"), whereas the expressive function is oriented toward the management of tension within the system ("integrative"). Unfortunately, this definition is neither unidimensional, nor unambiguous. For example, in his cross-cultural study of sex-role differentiation on this dimension, Zelditch (15) used multiple and not necessarily correlated criteria, such as boss vs. conciliator, emotional constraint vs. warmth, and discipline vs. permissiveness. Also, such behaviors as emotional constraint and discipline may often be used for the purpose of managing intrafamilial tensions. This dimension shall be defined here as extent of agreement between interactors concerning the explicit or implied goals of an interaction sequence. Behavior that functions to facilitate the other's act is taken to signify goal agreement, whereas that which functions to interfere with the other's act is taken to signify goal disagreement. For example, A's praise of B's action functions as a facilitation, whereas A's blame of B's action functions as an interference. In addition to its explicitness, this particular formulation of the

³ For purposes of simplification, the power and function dimensions are treated here as if they were dichotomous. Also, this study does not deal with the additional role types generated when older and younger siblings, adult relatives, etc., are considered.

⁴ Parsons distinguishes between two phases of this period, called oral dependency and love dependency. These two phases are treated as one in this study because the theory does not imply discontinuity between them with respect to role discriminations on the power dimension.

WALTER EMMERICH

instrumental-expressive dimension is similar to those used to distinguish sex roles in the child research literature (e.g., 5, 7, 13).

The classification of twelve kinds of social action according to the power and function dichotomies is given in Table 1. The specific items used in this study are found below each kind of action. It is assumed that each item is representative of the class of behaviors denoted by its corresponding kind of action. Application of the hypotheses discussed earlier to this scheme results in the following predictions: (a) With increasing age, young children will increasingly assign high power actions to parents and low power actions to children. (b) With increasing age, young children will increasingly assign goal agreement actions to female family members and goal disagreement actions to male family members. (c) The discrimination of age roles on the power dimension will occur at an earlier age than will the discrimination of sex roles on the function dimension.

TABLE I
CLASSIFICATION OF SOCIAL ACTS (ITEMS) ACCORDING TO POWER
AND GOAL AGREEMENT

GOAL AGREEMENT	P O W E R	
	High	Low
Agree (Facilitation)	1. <i>Gratification</i> You can have it.	7. <i>Asking</i> Can I have this?
	2. <i>Acceptance</i> Thank you for doing it.	8. <i>Conformity</i> I'll do as you say.
	3. <i>Praise</i> You did that very well.	9. <i>Adequacy</i> I can do it very well.
Disagree (Interference) .	4. <i>Deprivation</i> You can't have it.	10. <i>Demand</i> Give me that.
	5. <i>Control</i> Stop doing that.	11. <i>Deviation</i> No, I won't do it.
	6. <i>Blame</i> You did that wrong.	12. <i>Inadequacy</i> I can't do it very well.

Brim (1) has constructed a set of analytic categories drawn from role and social systems theory and applied these to studies of parent-child relationships. It is of interest that our explicit derivation of variables from the same theoretical framework leads to their clear-cut classification in Brim's terms: (a) This study assesses children's conceptions of actual role performances, rather than role prescriptions.⁵ (b) Each role has its reciprocal role

⁵ An important question for developmental social psychology concerns the ages when specific populations first respond differentially to the same stimuli under discrimination (cognitive) and prescriptive (evaluative) sets. If such a differential response has not yet developed in children of the age range sampled here, then the pattern of role discriminations found would be the same as that for role prescriptions.

CHILD DEVELOPMENT

included in the set of items. These reciprocals, such as asking-gratification and conformity-acceptance, are found opposite one another in the rows of Table 1. (c) The roles are described as overt behavioral units, rather than as motivational or effectual units. (d) The hypotheses concern role conceptions at the level of parent-child and female-male family systems. The hypotheses will be tested also at the lower systems level of dyadic relationships among the four basic family roles. This will be done in order to discover whether or not conceptualizations at the higher levels can be inferred directly from those at the lower levels.

In addition to the theoretically derived hypotheses concerning age changes, this study investigates sex differences in young children's discriminations of family roles. A variety of empirical studies have considered sex differences in children's perceptions of sex roles. Studies by Rabban (12) on appropriate sex-typing, Brown (2) on sex-role preference, and Emmerich (5) on sex-role identification, for example, suggest that young boys are more sensitive to sex-role differences than young girls. On the other hand, Kagan (7) found that a group of first and second graders perceive the father as less friendly, more punitive, and more dominant than the mother, and that girls make this discrimination more sharply at this age than boys. Brim's review (1) of several studies of parental child-rearing attitudes suggests that parents are more severe toward their same-sex children than toward their opposite-sex children. The apparent contradictions among these and other studies are probably due in large part to differences in their concepts, populations, specific measures, and methods of analysis. Comparison is particularly difficult because these studies deal with quite different sex-role processes for which the precise interrelations are unknown. As Brown (2, 3) points out, sex-role preference and identification may not be correlated responses in certain populations, and Lynn (9) makes essentially the same point about sex-role adoption and identification. The author would also suggest that sex-role discriminations are separate from and ordinarily precede any particular sex-role preference, identification, or adoption. In this regard, the present study was designed to find out if there are sex differences in young children's perceptions of sex roles.

Much less attention has been given to the question of sex differences in children's perceptions of age roles. In his interpretation of evidence for greater sex-role identification in boys than in girls, the writer speculated in an earlier study (5) that, relative to boys, girls identify more with those aspects of the mother that distinguish the mother and child roles than with those aspects that distinguish the mother and father roles. If this is the case, then it might be expected that parent and child roles will be discriminated more sharply by girls than by boys. Of course, this prediction is not a necessary implication of the above hypothesis, since boys and girls could identify differentially with the adult characteristics of parent roles while not differing in the extent to which these roles are discriminated.

WALTER EMMERICH

TABLE 2
SAMPLE CHARACTERISTICS

Item	Young Girls	Young Boys	Old Girls	Old Boys	Total Young	Total Old	Total Girls	Total Boys	Grand Total
N	22	22	22	22	44	44	44	44	88
Mean Age (months)	53.6	53.6	69.6	69.7	53.6	69.7	61.6	61.7	61.7
Age Range (months)	42-65	44-63	66-73	65-72	42-65	65-73	42-73	44-72	42-73
Age SD (months)	7.1	6.3	1.8	1.9	6.6	1.9	9.6	9.4	9.4
Socioeconomic Level (freq.)									
I & II	15	16	9	12	31	21	24	28	52
III & IV	6	4	5	7	10	12	11	11	22
College Student	1	2	8	3	3	11	9	5	14
Number of Siblings (freq.)									
None	1	2	1	0	3	1	2	2	4
One	5	13	10	10	18	20	15	23	38
Two	15	5	8	8	20	16	23	13	36
Three	1	2	3	3	3	6	4	5	9
Four	0	0	0	1	0	1	0	1	1
Examiner (freq.)									
Male A	5	6	0	1	11	1	5	7	12
Male B	7	4	7	11	11	18	14	15	29
Female	10	12	15	10	22	25	25	22	47

METHOD

Subjects

Subjects were white nursery school and kindergarten children from middle-class families unbroken by separation, divorce, or death. None was adopted. The four sex-age subsamples were formed from a larger original sample in such a way as to create nonoverlapping age ranges between "young" and "old" subgroups, the same mean age for both sex groups within each age level, and equal *N*s in all subgroups. The sample characteristics are given in Table 2. The mean age was 4 years, 5 months, for the young group, and 5 years, 10 months, for the old group. Over half of the total sample was clearly from socioeconomic levels 1 and 2 in Warner's (occupational) classification scheme (14), and over 80 per cent of the total sample had either one or two siblings.

Materials

Stylized figure drawings representing the four roles were prepared on 6¼ by 3¼ inch Hunter Card Master cards in the following pairs: mother-

CHILD DEVELOPMENT

girl, father-boy, mother-father, and girl-boy. In order to present each pair with all items in a nonconsecutive pattern, 12 cards were made of the same pair, with each member of the pair appearing on the left and right half the time. Female and male figures differed with respect to long vs. short hair, blouse and skirt vs. shirt and shorts, and triangular- vs. rectangular-shaped legs. Parent and child figures differed with respect to size, apron on the mother vs. no apron on the girl, and suit and necktie on the father vs. shirt and shorts on the boy. In order to add some variety to the stimuli, a third of the pairs were colored brown, a third blue, and a third red. All cards were sprayed with a coat of clear plastic to preserve the grease-pencil figures. The 48 cards were presented to subjects by means of the Hunter Card Master, Model 340.

Procedure

Subjects were tested individually after a brief period of group familiarization with *E*, the testing room, and the assessment materials. Subject sat in front of the exposure window of the Card Master, with *E* seated to his right. Experimenter pointed to the exposed mother-father card and said, "Do you see these people here? Let's play a game with them. Okay? Who says, 'I'm the mother'?" After *S* pointed to the correct figure, *E* said, "That's right," and then asked, "Who says, 'I'm the daddy'?" After *S* pointed to the correct figure, *E* exposed the next card which was the boy-girl pair, and said, "Who says, 'I'm the girl'?" After *S* pointed to the correct figure, *E* said, "Yes, that's a girl just like you," if *S* was a girl, or "Yes, that's a girl just your age," if *S* was a boy. Experimenter then asked, "Who says, 'I'm the boy'?" After *S* pointed to the correct figure, *E* said, "Yes, that's a boy just your age," if *S* was a girl, or "Yes, that's a boy just like you," if *S* was a boy. In order to have some practice with actual test items, the girl-mother pair was presented next with the item, "Who says, 'You can have it'?" and the father-boy pair was then presented with the item, "Who says, 'No, I won't do it'?" The test proper then began and consisted of presenting the 48 pairs in prearranged order, with *E* recording the figure chosen for each pair. Either pointing to a figure or naming it was an acceptable response. Subject was given as much time as he wanted, and *E* would repeat the item whenever *S* did not respond after several seconds or responded ambiguously. The average testing time was about 11 minutes with little variation among subsamples. Each child was told that he did very well at the completion of testing.

The sequence of presentation of paired figures and items is given in Table 3 and was the same for all subjects. This particular sequence was the result of the following procedures for reducing or controlling a variety of factors, particularly irrelevant response sets: each item was presented once within each quarter of the test; high and low power items were alternated in a nonsystematic sequence within each quarter, as were goal-agreement and goal-disagreement items; a specific pair of figures was never presented

WALTER EMMERICH

twice in succession; whenever possible, a specific pair of figures was not presented again until the other three pairs appeared in the sequence; when a specific pair of figures appeared again in the sequence, the positions of the figures were reversed; successive pairs were never of the same color; whenever possible, pairs in the other two colors were presented before a pair of the same color was repeated.

TABLE 3
SEQUENCE OF PRESENTATION OF PAIRED FIGURES AND ITEMS

Place in Sequence	Pair	Item Number	Place in Sequence	Pair	Item Number	Place in Sequence	Pair	Item Number
1	F-M	1	17	M-G	9	33	B-F	11
2	G-B	11	18	B-F	7	34	G-B	4
3	M-G	3	19	G-B	6	35	M-G	8
4	B-F	4	20	F-M	4	36	F-M	6
5	G-M	6	21	F-B	8	37	B-G	8
6	M-F	8	22	M-F	3	38	F-B	6
7	F-B	9	23	G-M	11	39	G-M	4
8	B-G	7	24	B-G	1	40	M-F	11
9	F-M	5	25	M-G	5	41	M-G	1
10	M-G	12	26	G-B	9	42	B-F	3
11	G-B	2	27	F-M	7	43	G-B	10
12	B-F	10	28	B-F	2	44	F-M	2
13	G-M	2	29	G-M	10	45	F-B	12
14	M-F	10	30	M-F	12	46	G-M	7
15	B-G	12	31	F-B	1	47	M-F	9
16	F-B	5	32	B-G	3	48	B-G	5

In order to have at least partial control over influences of examiner characteristics upon the children's responses, two males and one female served as experimenters. The number of subjects tested by each *E* in each subsample is given in Table 2.

Data Analysis

The major method of data analysis was that described by Edwards (4) for analysis of variance of repeated measures from several independent groups. Two kinds of levels \times subjects designs were used: $2(\text{ind.}) \times 2(\text{repeat})$, and $2(\text{ind.}) \times 2(\text{repeat}) \times 2(\text{repeat})$. The usual assumptions underlying the use of the *F* distribution apply to tests of significance between independent groups, but further assumptions described by Lindquist (8) are necessary for tests among repeated measure groups. In the analyses presented, inspection of subsample distributions and tests of homogeneity of variance did not give reason seriously to doubt that these assumptions were met.

CHILD DEVELOPMENT

Item Analysis

The internal consistency of each item was determined for the measure of extent of use of the power dimension to discriminate parent and child roles. Because a significant tendency was found for subjects to assign high power items to parents more than low power items to children (see Table 4), separate item analyses were conducted for the 12 cross-generation comparisons on low and high power items. Since each item was presented with the mother-girl and father-boy pairs, these were combined to form "one item" in the analyses. The phi coefficients and chi square tests of their significance were found for each item's ability to discriminate between the upper and lower thirds of the total sample. In all twelve analyses, χ^2 was significant ($p < .001$). The phis ranged from .52 to .78 with the median at .61, while the corresponding ϕ_{\max} coefficients ranged from .62 to .96 with the median at .75.

RESULTS AND DISCUSSION

Discrimination of Age Roles on the Power Dimension

The measure of extent of use of the power dimension to discriminate parent and child roles was the number of times the child assigned high power statements to the parent figures and low power statements to the child figures in the cross-generation comparisons. A score of 6 on this measure indicates that the child randomly assigned high and low power items to the parent and child figures, whereas a score of 12 indicates that the child consistently assigned high power items to parent figures and low power items to child figures. In order to explore the possibility that children might be more sensitive to either the high or low power items in discriminating age roles, the analysis was broken down so that power becomes an independent variable in the design. The means are given in Table 4 and the analysis of variance is summarized in Table 5.

The hypothesis that the older group would have a higher mean score on this measure than the younger group was confirmed ($p < .001$). Essen-

TABLE 4
GROUP MEANS ON USE OF POWER DIMENSION TO
DISCRIMINATE AGE ROLES
(Sexes Combined)

Item Type	AGE OF CHILD		
	Young	Old	Total
High Power	7.73	9.39	8.56
Low Power	5.57	7.39	6.48
Total	6.65	8.39	7.52

WALTER EMMERICH

TABLE 5

SUMMARY OF ANALYSIS OF VARIANCE FOR USE OF POWER DIMENSION
TO DISCRIMINATE AGE ROLES
(Sexes Combined)

Source of Variation	<i>df</i> ₁	<i>df</i> ₂	F	p
(A) Age of Child	1	86	20.85	< .001
(P) Power (High vs. Low Power Items)	1	86	17.16	< .001
A × P	1	86	.02

tially the same age difference was found for the mother-girl and father-boy comparisons when these were analyzed separately, indicating that the discrimination of parent and child roles on this dimension is derivable from discriminations at the lower level, mother-girl and father-boy social systems.

In an analysis of variance not included here, sex of the child was found to have no significant effect upon this measure, either singly or in interaction with the other independent variables. Thus, the hypothesis that girls would be more sensitive than boys to age role discriminations was not confirmed. Consequently, if a sex difference is found in future studies of young children's identifications with the age-role characteristics of their parents, such differences will probably not be due to a sex difference in the discrimination of age roles on the power dimension.

There was a significant tendency to assign high power items to parent figures more than low power items to child figures ($p < .001$). This result is not an artifact of differences in content between high and low power items, as low power items were no less homogeneous than high power items. This finding means that young children discriminate parent and child roles more on the basis of presence of high power in the parent and its absence in the child than on the basis of absence of low power in the parent and its presence in the child. Apparently, the child tends to define his role negatively rather than reciprocally; i.e., he perceives his lack of parental power but does not clearly conceive of his role as complementary and subordinate to that of the parent. Several reasons can be advanced to account for this finding: (a) Because high power acts of the parents are directly observed by the child, the parent role is rather easily discriminated, whereas the appropriateness of low power action for the child role must be learned through a slower process of trial and error. Although this explanation has appeal because of its simplicity, it is not supported by the finding of little or no decrease in discrepancy between high and low power item scores with increasing age. (b) The prescribed roles of higher status members of a social system such as the family are more clearly perceived and/or accepted by all members of the system than are the prescribed roles of lower status members. (c) Persons of high status more clearly perceive and/or accept

CHILD DEVELOPMENT

their prescribed roles than do persons of low status. (d) There is more toleration of deviation in the performance of low status roles than of high status roles. More detailed research on the discrimination and acceptance of age roles by older children and parents is necessary before any of these or other explanations can be accepted.

Discrimination of Sex Roles on the Function Dimension

The measure of extent of use of the function dimension to discriminate female and male roles was the number of times the female figure was chosen in the cross-sex comparisons. Scores for the facilitation and interference items were entered separately in the analysis so that function was an independent variable in the design. Because of the finding that children discriminate age roles on the power dimension, the above measure was derived from the mother-father comparisons on high power items only and from the girl-boy comparisons on low power items only. A score of 0 (3) means that the male (female) was chosen in every comparison. Thus, a group mean near 1.5 indicates that male and female figures were chosen about equally often. The means are given in Table 6 and the analysis of variance for each sex is summarized in Table 7.

The hypothesis that female figures would be assigned the facilitation function and male figures the interfering function was confirmed in girls ($p < .005$). In boys, however, a different and surprising pattern emerged, as seen from the significant interaction between generation and function ($p < .05$). Whereas boys also perceived the mother as more facilitating and the father as more interfering, the roles were reversed for the child-figures, as the *boy* was perceived as more facilitating and the *girl* as more interfering. This finding has important methodological and theoretical implications. Heretofore, it has been generally assumed that sex-role discriminations

TABLE 6
GROUP MEANS ON USE OF THE FUNCTION DIMENSION
TO DISCRIMINATE SEX ROLES

Item Type	G I R L S			B O Y S		
	Young	Old	Total	Young	Old	Total
M-F, Facilitation	1.91	2.05	1.98	1.68	1.73	1.70
M-F, Interference	1.73	1.68	1.71	1.18	1.23	1.21
G-B, Facilitation	1.77	2.23	2.00	1.41	1.36	1.39
G-B, Interference	1.32	1.45	1.39	1.73	1.45	1.59
Total M-F	1.82	1.86	1.84	1.43	1.48	1.45
Total G-B	1.55	1.84	1.69	1.57	1.41	1.49
Total Facilitation	1.84	2.14	1.99	1.55	1.55	1.55
Total Interference	1.52	1.57	1.55	1.45	1.34	1.40
Total	1.68	1.85	1.77	1.50	1.44	1.47

WALTER EMMERICH

TABLE 7

SUMMARY OF ANALYSES OF VARIANCE FOR USE OF FUNCTION
DIMENSION TO DISCRIMINATE SEX ROLES

Source of Variation	<i>df</i> ₁	<i>df</i> ₂	G I R L S		B O Y S	
			<i>F</i>	<i>p</i>	<i>F</i>	<i>p</i>
(A) Age of Child	1	42	1.3311
(G) Generation (M-F vs. G-B comparisons)	1	126	1.1906
(F) Function (Facilitation vs. inter- ference items)	1	126	10.67	< .005	1.13
A × G	1	126	.8555
A × F	1	126	.8518
G × F	1	126	1.58	6.42	< .05
A × G × F	1	126	.0615

are sufficiently generalized to be pretty much the same no matter what other characteristics happen to be associated with the particular opposite-sex objects being compared. Here, however, the age of the opposite-sex pair compared was a critical factor in how boys discriminate the sex roles. This result points up the importance of basing measures of sex-role discrimination upon representative or stratified samples of pairs of opposite-sex persons. In terms of social systems analysis, this finding indicates that, for the population of boys sampled, conceptualization of the female and male roles on the function dimension cannot be derived from sex discriminations at the lower level mother-father and girl-boy systems.

The correct explanation of this sex difference awaits further study of differential sex-role processes in the two sexes. One possible interpretation emphasizes certain motivational aspects of the assessment procedure. It seems reasonable to suppose that the figure of the subject's sex and age symbolizes the subject himself. When the boy is asked to indicate whether a figure representing himself is more or less positive than a figure representing a girl of his age, he might assign the positive actions to himself in order to avoid whatever anxiety is associated with negative self-attribution. This interpretation implies that the boy's perception of his sex role involves a systematic distortion of the role that he actually performs and would seem to be supported by the frequent observation that boys are more aggressive than girls.

It might be argued that this result does not actually reflect a stable characteristic of boys but is purely situational in the sense that boys were simply trying to impress the examiner with how "nice" they are, perhaps in hopes of securing the examiner's nurturance. However, if this had been the case, one might have also expected the boys interviewed by the woman

CHILD DEVELOPMENT

to allocate positive actions to the mother and those interviewed by the men to allocate positive actions to the father. Such an effect, if present, could not have been very strong because half the boys were tested by examiners of each sex and yet boys as a group clearly perceived the mother as more facilitating and the father as more interfering. Further, even if boys were trying to impress the examiner with how "nice" they were, such an act might actually reflect a generalized personality characteristic that transcends this particular assessment situation.

However, another question can be raised with regard to the above interpretation. Although it has not yet been adequately demonstrated, there is reason to believe that parents expect their young children to take sex-typed roles (3, 6). But if this is the case, why would boys need to avoid attributing characteristics to themselves which, although negative, are prescribed as part of their role? Indeed, there is evidence that boys are especially encouraged and pressured by socializers to adopt the masculine role (e.g., 2, 5, 6, 9). Of course, it is quite possible that parents expect boys to take on aspects of the father's sex role *other than* his interference behavior, such as the father's masculine skills, interests, and values.

The writer would venture a different and frankly speculative interpretation which accepts at face value the conclusions that girls conceptualize sex roles on the function dimension at a level that is sufficiently generalized to apply to both generations, whereas boys apply different sex-role conceptions to each generation. As noted above, previous research makes it quite reasonable to assume that the rewards and pressures upon the young boy to identify with and/or adopt the adult's sex role are at least as great if not greater than those upon the young girl. These conclusions suggest possible sex differences in the processes of sex-role identification. The fact that the girl's perception of her sex role on the function dimension is essentially the same as her perception of the mother's sex role on this dimension does not necessarily imply that the girl *identifies* with the mother's facilitating behavior, if by identification one means perceived or fantasied similarity. Rather, it would be more accurate to say that the girl perceives essentially *no difference* between herself and her mother on this dimension, relative to males. Perhaps this is identification in the sense that the girl places herself and the mother at about the same point on the function dimension, but, since it does not presuppose that the mother and girl are even discriminated in their sex roles, it might best be called an *undifferentiated* identification. Consequently, it does not seem necessary to postulate a more active identification process in the girl to account for her perception of herself as having or adopting the mother's facilitating behavior, although such a process may have occurred at an earlier period in the girl's development. The boy, however, assigns interfering behavior to his father and facilitating behavior to himself, relative to females. If he is to identify with the father's sex role on this dimension, he must do so in spite of this difference in their sex-role definitions. He may feel compelled to identify with the

WALTER EMMERICH

father on the very dimension that defines the father and son roles at opposite poles. In order to succeed, such an identification process might require an active integration of the two roles into what can be called, in contrast to the girl's case, a *differentiated* identification. Perhaps the exaggerated, compensatory quality of the masculine behavior that has been observed in young boys (9) is a manifestation of the boy's attempts to identify with the father's sex role under these conditions. The writer plans to explore these hypotheses in a future study.

The hypothesis that older children would make sharper sex-role discriminations than younger children was not confirmed for either sex. Further, there was no support for Parsons' contention that age-role discriminations occur earlier developmentally than sex-role discriminations. However, before rejecting Parsons' sequential hypothesis altogether, it would be important to demonstrate empirically that the findings of this study can be linearly extended to adjacent age periods and that age roles are not discriminated at an earlier age on a different dimension.

Discrimination of Age Roles on the Function Dimension

The results presented thus far support Parsons' contention that young children discriminate age roles on the power dimension and sex roles on the function dimension. The theory also implies that young children define family roles primarily in terms of each member's intrafamilial relationships, rather than on the basis of one's participation in other subsystems of the society. For example, young children are probably much less likely to describe the father in terms of a breadwinning role than in terms of an interfering role. It is believed that a correlate of this kind of structural simplicity is the use of one and only one dimension to make each role discrimination. Specifically, it would be predicted that young children do *not* discriminate age roles according to function nor sex roles according to power. This and the following section are concerned with testing these hypotheses.

The measure of extent of use of the function dimension to discriminate age roles was the number of times the child assigned facilitation items to adult figures and interference items to child figures in the cross-generation comparisons. The means are given in Table 8. The hypothesis was tested by the significance of the difference between the means and the score of 12.00, which signifies random assignment of facilitating and interfering items to parent and child roles. In boys, this difference is not significant when tested separately for young and old age groups but does reach significance when the age groups are combined ($t = 2.03, p < .05$). In girls, the difference for the old group only is significant ($t = 3.53, p < .01$). The direction of the significant differences indicates that children tend to assign interfering behavior to the adult role and facilitating behavior to the child role. The magnitudes of these differences do not seem great enough to warrant more detailed interpretation and are clearly not as great as those

CHILD DEVELOPMENT

TABLE 8

GROUP MEANS ON USE OF THE FUNCTION DIMENSION TO
DISCRIMINATE AGE ROLES

	A G E O F C H I L D		
	Young	Old	Total
Girls	12.27	10.59	11.43
Boys	11.14	11.45	11.30
Total	11.71	11.02	11.37

found for the use of the *power* dimension to discriminate age roles. Nevertheless, these findings do indicate that children of this age may use the function dimension to discriminate age roles.

Discrimination of Sex Roles on the Power Dimension

The measure of extent of use of the power dimension to discriminate female and male roles was the number of times the female figure was chosen in the cross-sex comparisons. Scores for high and low power items were entered separately in the analysis so that power was an independent variable in the design. In order to explore the possibility of greater sensitivity to this dimension within the adult or child generation, scores for mother-father and girl-boy comparisons were also entered separately in the analysis. The means are given in Table 9 and the analysis of variance for each sex is summarized in Table 10.

TABLE 9

GROUP MEANS ON USE OF THE POWER DIMENSION
TO DISCRIMINATE SEX ROLES

Item Type	G I R L S			B O Y S		
	Young	Old	Total	Young	Old	Total
M-F, High Power	3.73	3.73	3.73	2.86	2.95	2.91
M-F, Low Power	3.05	2.95	3.00	2.82	2.73	2.78
G-B, High Power	2.86	3.09	2.97	3.05	2.27	2.66
G-B, Low Power	3.23	3.68	3.46	3.14	2.77	2.96
Total M-F	3.39	3.34	3.36	2.84	2.84	2.84
Total G-B	3.05	3.38	3.22	3.10	2.52	2.81
Total High	3.30	3.41	3.35	2.96	2.61	2.78
Total Low	3.14	3.32	3.23	2.98	2.75	2.86
Total	3.22	3.36	3.29	2.97	2.68	2.82

WALTER EMMERICH

TABLE 10

SUMMARY OF ANALYSES OF VARIANCE FOR USE OF THE POWER
DIMENSION TO DISCRIMINATE SEX ROLES

Source of Variation	df_1	df_2	G I R L S		B O Y S	
			F	p	F	p
(A) Age of Child	1	42	.36	1.02
(G) Generation (M-F vs. G-B comparisons)	1	126	.4204
(P) Power (High vs. low items)	1	126	.3018
A \times G	1	126	.8804
A \times P	1	126	.03	1.61
G \times P	1	126	7.10	<.01	1.33
A \times G \times P	1	126	.0040

Boys did not consistently use the power dimension to discriminate sex roles. Girls, however, assigned the more powerful role to the *mother* (relative to the father) and to the *boy* (relative to the girl) ($p < .01$). Thus, the hypothesis that the power dimension is not used to discriminate sex roles seems to hold for boys but not for girls. The pattern in girls is similar to that found for boys' use of the function dimension to discriminate sex roles in that the direction of the discrimination depends upon the generation within which the sexes are compared. In terms of social systems analysis, this finding indicates that, for the population of girls sampled, conceptualization of the female and male roles on the power dimension cannot be derived from sex discriminations at the lower level mother-father and girl-boy systems.

These results raise further questions about sex-role perception and acquisition in girls. In view of the commonly held but not fully substantiated belief that greater power is attached to the adult familial role of males than females in our society, it is surprising that girls perceived the mother as relatively more powerful than the father. Whether this represents a systematic distortion or a realistic appraisal of the family social structure experienced by the young girl remains to be determined. In either case, however, this situation may present the girl with a problem of sex-role identification and/or adoption that is analogous to that discussed earlier in connection with the boy's identification with the father's sex role on the function dimension. Insofar as the girl feels compelled to identify with the mother on the power dimension, she must do so despite her perception of herself as different from the mother with respect to power in *both* her age- and sex-role characteristics. It may require, then, an active *differentiated* identification process in order for the girl to integrate her role with the

CHILD DEVELOPMENT

mother's on the power dimension. This hypothesis will also be explored in a future study.

SUMMARY

This study investigated some implications of Parsons' theory concerning the development of parent and child role concepts in young children. A modified paired-comparison procedure was used to assess the extent to which "young" (mean CA, 4-5) and "old" (mean CA, 5-10) groups of middle-class children used two dimensions to discriminate age and sex roles. These dimensions were function (facilitation vs. interference), a modification of Parsons' instrumental-expressive distinction, and power (high vs. low). The major findings were:

1. With increasing age, there is increased use of the power dimension to discriminate age roles.
2. Children perceive their lack of parental power but less clearly conceive of the child role as complementary and subordinate to the parent role.
3. Facilitating behavior is allocated more to the mother's sex role and interfering behavior more to the father's sex role.
4. Children assign facilitating behavior more to themselves and interfering behavior more to their opposite-sex peers.
5. Girls perceive the mother as more powerful than the father but see themselves as less powerful than the boy.

The results were discussed in social role and systems terms and were found to support some aspect of Parsons' theory and to suggest modifications of others. They were also interpreted as indicating the need for: (a) basing measures of sex-role discrimination upon representative or stratified samples of opposite-sex pairs; (b) a more precise understanding of the actual family social structure impinging upon the young child; and (c) exploration of possible sex differences in identification processes arising from sex differences in sex-role discriminations.

REFERENCES

1. BRIM, O. G., JR. The parent-child relation as a social system: I. Parent and child roles. *Child Developm.*, 1957, 28, 343-364.
2. BROWN, D. G. Sex-role preference in young children. *Psychol. Monogr.*, 1956, 70, No. 14 (Whole No. 421).
3. BROWN, D. G. Sex-role development in a changing culture. *Psychol. Bull.*, 1958, 55, 232-242.
4. EDWARDS, A. L. *Experimental design in psychological research*. New York: Rinehart, 1950.
5. EMMERICH, W. Parental identification in young children. *Genet. Psychol. Monogr.*, in press.

WALTER EMMERICH

6. FAULS, LYDIA B., & SMITH, W. D. Sex-role learning of five-year-olds. *J. genet. Psychol.*, 1956, 89, 105-117.
7. KAGAN, J. The child's perception of the parent. *J. abnorm. soc. Psychol.*, 1956, 53, 257-258.
8. LINDQUIST, E. F. *Design and analysis of experiments in psychology and education*. Boston: Houghton-Mifflin, 1953.
9. LYNN, D. B. A note on sex differences in the development of masculine and feminine identification. *Psychol. Rev.*, 1959, 66, 126-135.
10. PARSONS, T. Family structure and the socialization of the child. In T. Parsons, R. F. Bales & Associates. *Family, socialization, and interaction process*. Glencoe, Ill.: Free Press, 1955. Pp. 35-131.
11. PARSONS, T., & SHILS, E. A. (Eds.) *Toward a general theory of action*. Cambridge: Harvard Univer. Press, 1951.
12. RABBAN, M. Sex-role identification in young children in two diverse social groups. *Genet. Psychol. Monogr.*, 1950, 42, 81-158.
13. SEARS, R. R., MACCOBY, ELEANOR E., & LEVIN, H. *Patterns of child rearing*. Evanston, Ill.: Row, Peterson, 1957.
14. WARNER, W. L., MEEKER, MARCHIA, & EELLS, K. *Social class in America*. Chicago: Science Research Associates, 1949.
15. ZELDITCH, M., JR. Role differentiation in the nuclear family: a comparative study. In T. Parsons, R. F. Bales & Associates. *Family, socialization, and interaction process*. Glencoe, Ill.: Free Press, 1955. Pp. 307-351.

THE UNIVERSITY OF CHICAGO

THE DIVISION OF THE PHYSICAL SCIENCES

THE DIVISION OF THE CHEMICAL SCIENCES

THE DIVISION OF THE BIOLOGICAL SCIENCES

THE DIVISION OF THE SOCIAL SCIENCES

THE DIVISION OF THE HUMANITIES

THE DIVISION OF THE ARCHITECTURAL SCIENCES

THE DIVISION OF THE ENGINEERING SCIENCES

THE DIVISION OF THE MEDICAL SCIENCES

THE DIVISION OF THE AGRICULTURAL SCIENCES

THE DIVISION OF THE ENVIRONMENTAL SCIENCES

THE DIVISION OF THE EARTH SCIENCES

THE DIVISION OF THE PLANETARY SCIENCES

THE DIVISION OF THE COSMOSCIENCES

THE DIVISION OF THE INTERDISCIPLINARY SCIENCES

THE DIVISION OF THE EDUCATIONAL SCIENCES

THE DIVISION OF THE LIBERAL ARTS SCIENCES

THE DIVISION OF THE PROFESSIONAL SCIENCES

THE DIVISION OF THE APPLIED SCIENCES

THE DIVISION OF THE TECHNICAL SCIENCES

THE DIVISION OF THE INDUSTRIAL SCIENCES

AN EVALUATION OF THE HIGHBERGER EARLY-ADJUSTMENT-TO-SCHOOL SCALE

WILLARD W. HARTUP*

State University of Iowa

Highberger (4) has described a method for measuring the young child's initial adjustment to nursery school. This Early-Adjustment-to-School Scale (EAS) is a global measure of adjustment composed of ratings on 54 behavior items. The items included in this scale cover a broad range of social and emotional behaviors commonly observed in groups of 2- and 3-year-old children. Most prominent in the scale would seem to be items relating to dependency-withdrawal, sociability, assertiveness, and some forms of aggression.

The procedure for the EAS includes: (a) systematic observation of the children in a preschool group from the fifth through fifteenth days of school attendance; (b) recording, during this time, the frequency with which the individual child evidences each type of behavior included in the scale; (c) preparation of ratings based on the observation records; (d) combining the ratings into a total EAS score.

There are a number of uses for a reasonably objective method of evaluating the child's early adjustment to school. Workers with groups of children are often concerned about the effects of group experience on the 2- or 3-year-old. Yet teachers have lacked objective criteria with which to assess the child's behavior during the early weeks of preschool experience. Developmental psychologists have been interested in the young child's reactions to separation from his mother, both as a consequence of the nature of early mother-child interaction and as a predictor of later trends in personality development. The Highberger technique offers promising aid in approaching problem areas such as these. Although it is based on a broad definition of adjustment (and thus poses some problems for the researcher or teacher), the EAS introduces greater precision to the measurement of initial school adjustment than has been the case heretofore. The EAS is, however, extremely cumbersome and time-consuming (hence expensive) to use. Ten days of concentrated observation are needed prior to rating and the ratings themselves are difficult for an unsophisticated observer.

In her original publication Highberger reported data relating to the reliability of total EAS scores, the reliability of individual items, and the relation of EAS total scores to teacher judgments of the child's adjustment and to a measure of maternal behavior. It seems, however, that empirical

* Iowa Child Welfare Research Station, State University of Iowa, Iowa City.

CHILD DEVELOPMENT

evaluation of this scale is needed in addition to the material supplied by Highberger before wider use is made of this particular technique. Such an evaluation is the general purpose of this study.

Answers to two specific questions have been sought: (a) Of what use is the total EAS score as a predictor of the child's behavior after the initial adjustment period is past and after a considerable period of organized preschool experience? (b) To what extent is early school adjustment, as measured by EAS total scores, a consequence of parent-child interaction?

Two considerations guided the present exploration of the predictive capacity of the EAS: first, that EAS total scores be related to measures of child behavior secured a considerable length of time after the administration of EAS itself; second, that EAS scores be related to measures of child behavior of greater specificity than "adjustment." Such exploration should help to clarify the meaning of the total EAS score. Aspects of children's behavior measured in the present study included dependence, aggression, motor activity level, sociability, nurturant behavior, sociometric status, and participation in creative activities. All measures were obtained approximately seven months after the EAS ratings.

Highberger found no significant relationship between EAS scores and selected items from the Fels Parent Behavior Ratings. The present investigation extends this area to evaluation by relating EAS scores to two other measures of parent behavior: the Parental Attitude Research Instrument (11) and selected portions of the maternal interview developed by Sears, Maccoby, and Levin (12).

METHOD

Subjects

Thirty-three children in two school groups at the Preschool Laboratories of the Iowa Child Welfare Research Station were the subjects for this investigation. Group A, seven boys and eight girls, was studied during 1956-1957. Group B, 10 boys and eight girls, was studied during 1957-1958. None of the children had attended nursery school prior to this time. At entrance the children's ages ranged from 2-6 to 3-4 in Group A, from 2-6 to 3-3 in Group B. The mean age was 2-11 in both groups. The children came primarily from business and professional homes; the sample is therefore skewed upward in terms of socioeconomic status and probably also in intelligence.

The Early-Adjustment-to-School Scale

Highberger's observation and rating procedures for the EAS were followed as closely as possible in this investigation. Observations and ratings were made by two observers in Group A¹, by one of the two observers in Group B. The product-moment correlation between the total EAS scores of the two observer-raters in Group A was .92.

¹ The author is grateful to Mrs. Frances Horowitz for this assistance.

WILLARD W. HARTUP

TABLE 1

RANGES, MEANS, AND STANDARD DEVIATIONS OF EARLY-ADJUSTMENT-TO SCHOOL SCORES FOR TWO PRESCHOOL GROUPS

	<i>N</i>	<i>Range</i>	<i>Mean</i>	<i>SD</i>
<i>Group A</i>				
Boys	7	170-195	179.0	11.88
Girls	8	156-211	176.4	16.62
Total	15	156-211	177.6	14.66
<i>Group B</i>				
Boys	10	151-202	185.6	15.10
Girls	8	150-200	173.5	17.45
Total	18	150-202	180.2	17.27

The EAS is constructed so that high scores indicate relatively "good adjustment" while low scores indicate relatively "poor adjustment." Table 1 presents the ranges, means, and standard deviations of early adjustment scores for the boys and girls in the two preschool groups. All distributions appeared, by inspection, to be normal. The sex differences are not significant. Comparison of the means for Group A and Group B is not meaningful since each child's EAS score is a measure of his behavior relative to the behavior of other children in his preschool group.

Teachers' Ratings

Approximately seven months following the observations for the EAS, two teachers in each preschool group² rated the children on a series of seven-point scales regarding dependence, aggression, motor activity, and sociability. Each scale ranged from low to high frequency of behavior. The dependency ratings completed were: *seeks approval*, *seeks unnecessary help*, *seeks physical contact*, *seeks to be near*, *seeks positive attention*, and *seeks negative attention*. Aggression ratings included: *interpersonal physical aggression*, *interpersonal verbal aggression*, and *aggression toward inanimate objects*. One scale each was used for *motor activity* and *sociability*.

The child's frequency of participation in certain "creative" art activities was also rated by the teachers, but with four-point scales. Activities covered in this way were: *plastics* (dough, clay, etc.), *easel painting*, *finger painting*, *collage*, *drawing*, and *manipulative activities* (puzzles, etc.).

Each pair of teachers completed their ratings independently. Product-moment correlations between the ratings of the two teachers in Group A

² The author extends thanks to Miss Nancy Mann, Mrs. Nina Mogar, Miss Rosalind Gold, and Miss Helen Tieszen for this assistance. Miss Mary Alice Russell also contributed to this part of the data collection.

CHILD DEVELOPMENT

ranged between .73 and .92 with a median correlation of .85. Reliability coefficients in Group B were somewhat lower: these ranged between .50 and .88 with a median of .68.³ The child's score on a given scale is the sum of the two teachers' ratings.

Observations

Observational data regarding dependency and nurturance were collected for Group B in the course of other investigations (5, 8). Nurturance was defined as attempts by the child to extend help, affection, sympathy, and reassurance to others. Each child was observed for 55 two-minute intervals during a three-week period approximately six months after the observation for the EAS. Instances of dependence recorded during the observation intervals were classified according to the following categories: *seeks positive attention*, *seeks reassurance*, *seeks physical proximity*, *seeks physical affection*, and *seeks help*. Nurturance categories included: *gives affection*, *gives positive attention*, *gives reassurance*, and *gives protection*.

To establish observer reliability two observers made 250 two-minute observations of 10 children in another preschool group. Recording of behavior was done independently by the observers. Agreement for both the dependency and nurturance categories was 90 per cent. Frequencies in the dependency categories were distributed in such a way that it was possible to use subscores in the analysis of data. Frequencies in the nurturance categories were low, thus necessitating the use of summary scores in the analysis.

Sociometric Status

Moore (10) obtained popularity scores for the children in Group B. Her technique combined the picture-sociometric method (7) and a method for eliciting negative choices (2). Three scores for each child were obtained in this fashion: (a) positive choices—an index of how much the child was "liked" by the other children in the group; (b) negative choices—an index of how much the child was "disliked" by the other children; (c) total popularity—the child's positive choice score minus his negative choice score. The product-moment correlations between two administrations of the sociometric, one month apart, were .66 for the positive choices, .61 for negative choices, and .62 for total popularity scores.

The Parental Attitude Research Instrument

The PARI, described by Schaefer and Bell (11), was mailed to the mothers of Group A subjects for completion at home. A slightly modified form of the PARI (1) was administered to the mothers of children in Group B. These mothers met in small groups to complete the questionnaire.⁴

³ A complete description of these rating scales and the training procedures employed with the raters may be obtained from the author.

⁴ The author wishes to acknowledge the contributions of Dr. David H. Crowell to this part of the study.

WILLARD W. HARTUP

Three scales appearing on Schaefer and Bell's inventory were deleted in the modified form used for Group B mothers, and several other scales were reworded so that items were appropriate for both fathers and mothers.

The Maternal Interview

Interviews were conducted with each mother of the children in Group B. A modified version of the Sears *et al.* interview schedule (12) was used. The schedule employed for this study included only those questions which related to the 44 rating scales factor analyzed by Milton (9). The interviews, conducted in a room at the preschool by two members of the Iowa Child Welfare Research Station staff,⁵ were recorded by Ediphone. Typescripts of each interview were prepared and rated following the procedures detailed by Sears *et al.*

Pearson correlation coefficients were computed between the EAS and the various measures of child and maternal behavior described in this section. Correlations were computed separately for Group A and Group B since data were obtained from different raters at different times. In order to estimate the relationships between these measures for the entire population of children studied, the separate correlations were combined using Fisher's procedure (3).

RESULTS AND DISCUSSION

The EAS and Dependency

The correlations between EAS scores and the measures of dependency are presented in Table 2. Most of these correlations are inverse and of low order. Only one classification of dependency behavior, *seeking to be near*, is inversely related to the EAS at a level significantly different from zero. The negative correlation between the teachers' rating of *seeking to be near* and the EAS is significant for Group A and for the combined samples. The observation measure of *seeking to be near*, available for Group B only, is also inversely correlated with the EAS at a significant level. This consistency in the findings for both ratings and observation of *seeking to be near* suggests that this one relationship is probably not chance determined.

Although the correlations between most of the dependency categories and the EAS are inverse, the teachers' rating of *negative attention seeking* is positively correlated with the EAS. *Negative attention*, as defined in the teachers' rating, includes hostile, destructive, noncompliant behavior used by the child to get and maintain the attention of others. Defined in this way, *negative attention* is probably a confounded measure of both dependency and aggression. Since other forms of aggression are positively related to the EAS (see below), it seems as if *negative attention* resembles aggression in these data much more than it resembles dependence.

⁵ Appreciation is extended to Dr. Ruth Updegraff and Miss Rosalind Gold for use of this material.

CHILD DEVELOPMENT

TABLE 2

CORRELATIONS BETWEEN EARLY-ADJUSTMENT-TO-SCHOOL SCORES
AND MEASURES OF DEPENDENCY

<i>Dependency Measure</i>	<i>Group A (N=15)</i>	<i>Group B (N=18)</i>	<i>Combined</i>
<i>Teacher Ratings</i>			
Seek Approval20	.00	.09
Seek Unnecessary Help07	-.12	-.04
Seek Physical Contact	-.32	-.20	-.25
Seek to Be Near	-.57*	-.27	-.41*
Seek Positive Attention	-.35	-.14	-.23
Seek Negative Attention66**	.28	.46**
<i>Observations</i>			
Seek Reassurance		-.19	
Seek Help		-.19	
Seek Physical Contact		-.13	
Seek Physical Proximity		-.62**	
Seek Positive Attention		-.21	

* Significant beyond .05 level.

** Significant beyond .01 level.

Although 18 of the 54 EAS items concern dependence, the total score would be, at best, a confounded measure of this particular variable. It is not surprising, therefore, that most of the correlations in Table 2 are low. Even so, these data help to clarify the meaning of the total EAS score. There are indications that this score is negatively related to passive, withdrawal types of dependence and positively related to assertive, aggressive dependence.

TABLE 3

CORRELATIONS BETWEEN EARLY-ADJUSTMENT-TO-SCHOOL SCORES
AND TEACHERS' RATINGS OF AGGRESSION AND ACTIVITY LEVEL

	<i>Group A (N=15)</i>	<i>Group B (N=18)</i>	<i>Combined</i>
<i>Aggression</i>			
Physical Aggression67**	.43	.55**
Verbal Aggression73**	.35	.55**
Aggression to Equipment71**	.25	.48**
<i>Activity Level</i>			
Motor Activity70**	.62**	.66**
Social Interaction45	.67**	.57**

* Significant beyond .05 level.

** Significant beyond .01 level.

WILLARD W. HARTUP

The EAS and Aggression

Table 3 shows the correlations between the EAS and the teachers' ratings of aggression. These correlations are all positive and, in Group A, quite high. Children who obtained high scores on the EAS were, as compared with low scorers on that instrument, more physically aggressive to other persons in the nursery school, more verbally aggressive, and more aggressive toward equipment and furnishings. The positive correlations between these measures for Group B fail to reach significance. The correlations are, therefore, moderate in size for the combined samples, but remain significantly greater than zero.

The difference in size of the correlations in the two preschool groups may be a function of the reliability of the teachers' ratings of aggression. In Group A the child's aggressive behavior was measured by the pooled ratings of two teachers whose individual ratings were in high agreement. This was not the case for group B (see above).

TABLE 4

PARTIAL CORRELATIONS BETWEEN EARLY-ADJUSTMENT-TO-SCHOOL
SCORES AND THREE MEASURES OF AGGRESSION WITH
TWO MEASURES OF ACTIVITY LEVEL HELD CONSTANT

	Group A (N=15)	Group B (N=18)
<i>Partialling out Motor Activity Level</i>		
Physical Aggression29	.10
Verbal Aggression51*	-.03
Aggression to Equipment35	-.06
<i>Partialling out Social Interaction</i>		
Physical Aggression63*	.04
Verbal Aggression72**	-.03
Aggression to Equipment65**	-.10

* Significant beyond .05 level.

** Significant beyond .01 level.

The EAS contains only six items dealing with aggression. There are, however, 12 items in the EAS which relate to the child's activity level. It is possible that the positive correlations between the EAS and aggression might be functions of a correlation between aggression and activity level. In order to test this hypothesis the correlations between the EAS and aggression were computed with the effects of activity level partialled out. The teachers had rated the children on two types of activity level: physical mobility and frequency of social interaction. The partial correlations reported in Table 4

CHILD DEVELOPMENT

show that only verbal aggression remains significantly correlated with the EAS when motor activity, as measured by teachers' ratings, is held constant. The magnitude of the correlations between the EAS and aggression is virtually unaffected when sociability is partialled out.

These data suggest that there is only a tenuous independent relationship between the child's initial adjustment to nursery school, as measured by the EAS, and his later physical aggressiveness. However, the frequency with which the child expresses verbal hostility is related to the total EAS score. If the EAS may be regarded as a measure of "emotional maturity," these data support the frequently-stated hypothesis that verbal aggression is a relatively "mature" form of social behavior.

The EAS and Activity Level

The correlations shown in Table 3 indicate that the EAS is positively correlated with the teachers' rating of motor activity level in both Group A and Group B. The EAS is also positively related to sociability. These correlations are among the highest found in this study and are the only ones of this magnitude found in both groups of subjects. The EAS score obtained during the first weeks of nursery school appears to be a good predictor of the child's frequency of participation in large muscle activity and social interaction during the terminal stages of the school year.

The EAS covers rather extensively those aspects of a child's behavior that could be called "activity level," "participation in activities," "outgoingness," etc. To some extent, then, the EAS total score and the teachers' ratings of activity level must be based on similarly defined variables. If so, the findings indicate that patterns of mobility, both physical and social, tend to remain stable over a year of the sort of experience available to children in the Iowa Preschool Laboratories.

The EAS and Nurturance

The observational measure of nurturance was obtained for Group B only. Nurturance included giving affection, attention, reassurance, and protection to other children and adults in the group. In addition to classifying the child's behavior in this manner, the observers also attempted to judge whether the child's nurturance was used (a) as a type of dependence, or (b) without expectation of social approval, affection, etc. Dichotomizing the child's nurturant behavior in this manner was accomplished with observer agreement of 85 per cent.

Dependent-nurturance was not significantly correlated with the EAS ($r = .09$). *Nondependent-nurturance*, however, was positively correlated with the early adjustment measure ($r = .47, p = .05$). This correlation is, of course, based on a small number of subjects and is only moderate in size. These data suggest, however, that the EAS total score is a useful predictor of the child's assertive, outgoing nurturance toward others, but is not related to nurturance used by the child to elicit affection from others.

WILLARD W. HARTUP

The EAS and Sociometric Status

None of the correlations between the EAS and the various sociometric scores obtained for Group B is significantly different from zero. Correlations between the EAS and the positive choice scores from the two administrations of the sociometric were .37 and .38. The negative choice scores were correlated $-.08$ and $-.31$ with the EAS. Total popularity scores were correlated .22 and .38 with EAS. There are interesting relationships in these low order findings: children who adjust to school best tend to be (a) most liked by their peers, (b) least disliked, and (c) most generally popular. These results with preschool children are consistent with a number of studies with older children in which positive relationships have been found between sociometric status and adjustment (e.g., 6, 13). However, replication of these findings with a larger, more heterogeneous sample is necessary before making conclusive statements concerning the relationship between the EAS total score and sociometric status.

TABLE 5

CORRELATIONS BETWEEN EARLY-ADJUSTMENT-TO-SCHOOL SCORES
AND TEACHERS' RATINGS OF CHILD'S PARTICIPATION
IN CREATIVE ACTIVITIES IN THE PRESCHOOL

	Group A (N=15)	Group B (N=18)	Combined
Plastics	-.39	-.19	-.28
Painting, Easel	-.41	-.41	-.41*
Painting, Finger	-.18	-.07	-.12
Collage	-.38	-.13	-.24
Drawing	-.20	-.08	-.12
Manipulative	-.37	-.43	-.40*

* Significant beyond .05 level.

The EAS and Participation in Creative Activities

In the Iowa Laboratory Preschools graphic and plastic art materials are almost continuously available to the children, almost always under free choice conditions. The frequency with which the children participated in these activities was measured by teachers' ratings. The correlations between these measures and the EAS are presented in Table 5; they are all inverse and of nonsignificant order. When the correlations for the separate preschool groups are combined, two types of creative participation are correlated significantly with the EAS: easel painting and manipulative activity.

These correlations, although not high, indicate that children who initially are poorly adjusted to the school situation become, by the end of the

CHILD DEVELOPMENT

year, the children who are most active in the school's graphic arts program. This could result from several underlying conditions. First, children having a difficult initial time at school are probably generally anxious about initiating social contacts. Many tend to seek activities where social pressures and conflicts are at a minimum. Creative activities are usually quiet and slow-paced and so possibly represent escape opportunities for some socially anxious children. Second, it is possible that children who are frustrated and anxious in their social interactions may seek creative activity as a means of releasing feelings. The present data could reflect either or both of these conditions. In any event, the findings indicate that the personal-social adjustment of the young child may be one of the determinants of participation in the graphic and plastic arts program of a preschool.

TABLE 6
CORRELATIONS BETWEEN EARLY-ADJUSTMENT-TO-SCHOOL SCORES
AND MEASURES OF MATERNAL BEHAVIOR

Measure of Maternal Behavior	Group A (N=15)	Group B (N=18)
<i>Parental Attitude Research Instrument</i>		
Factor I (Auth.-Control)04	-.05
Factor II (Hostility-Rejection)13	-.10
Factor III (Democratic Attitudes)	-.05	-.02
<i>Interview</i>		
Factor A (Permissiveness-strictness) ...		-.17
Factor B (General family adjustment) ..		.23
Factor C (Warmth of mother-ch. rel.) ..		.25
Factor D (Responsible ch. train. orient.)		.30
Factor E (Aggressive-punitiveness) ...		-.24

The EAS and the PARI

The PARI is composed of 23 five-item subscales covering a broad range of parental attitudes toward child-rearing and family life. Zuckerman (14) has recently factor analyzed the PARI. He reports the emergence of three fairly-well defined factors. These factors formed the basis for the present analysis of the relationship between the PARI and the EAS. Scales loading .60 or more on Factor I (Authoritarian-Control) were grouped by simple addition of scores in the present analysis; scales loading .50 or more on Factor II (Hostility-Rejection) were similarly grouped, as were those loading .40 or more on Factor III (Democratic Attitudes). The correlations between the EAS and these three cluster-scores are presented in Table 6. There appear to be no significant relationships between the EAS total score and any of the PARI clusters for either group of subjects studied.

WILLARD W. HARTUP

The EAS and the Maternal Interview

Interview cluster scores were derived by summing the ratings on scales loading more than .30 on each of five factors identified by Milton (9). As can be seen from Table 6, none of these cluster scores is significantly correlated with the EAS.

The present results parallel Highberger's: no relationships exist between the EAS and broad measures of parent attitudes and practices. The Fels Parent Behavior Ratings, the PARI, and the Sears *et al.* interview all have proved independent of the EAS. It is possible that relationships do exist between parental practices and aspects of the young child's initial adjustment to nursery school but that these relationships involve nonglobal variables. Certainly the measure of adjustment to school and the measures of parent behavior used in this study are all complex. That such measures are clearly unrelated suggests that future interest in the relationship between parent practices and the child's early adjustment to school must be directed toward a basic refinement of the measures used.

It should also be remembered that groups of parents and children, homogeneous with respect to socioeconomic class and intelligence, were studied by both Highberger and the present investigator. Further study of the relationship between parent practices and the child's early school adjustment should probably not be attempted unless more heterogeneous groups of parents and children are employed.

SUMMARY

The purpose of this investigation was to explore certain correlates of the Early-Adjustment-to-School Scale (4). Thirty-three children in two pre-school groups were observed and rated on the EAS during their second and third weeks in nursery school. These ratings were then correlated with the following measures: (a) dependency, aggression, motor activity, sociability, nurturance, participation in creative activities, and sociometric status as shown by the child approximately seven months after the administration of the EAS; (b) measures of maternal child-rearing attitudes and practices.

Significant positive correlations were obtained between the EAS and measures of negative attention seeking (dependence), physical aggression, verbal aggression, nonpersonal aggression, motor activity, sociability, and nurturance. Significant negative correlations were obtained between the EAS and measures of seeking to be near (dependence) and participation in creative activities. The total EAS score has, then, certain utility as a long-term predictor of children's behavior in group situations. Considered together, these results also point to aggressive assertiveness vs. dependency-withdrawal as one sort of dimension within the EAS.

The EAS proved unrelated to the measures of maternal child-rearing attitudes and practices. The failure of the EAS to reflect patterns of mother-

CHILD DEVELOPMENT

child interaction is possibly due to the global character of both the EAS and the particular measures of maternal practices employed in the present study.

REFERENCES

1. CHANTINY, J. G., LOVELL, L. L., & McCANDLESS, B. R. *Parent practices research scale*. Iowa Child Welfare Research Station, State Univer. of Iowa, 1956.
2. DUNNINGTON, MARGARET J. Investigation of areas of disagreement in sociometric measurement of preschool children. *Child Developm.*, 1957, 28, 93-102.
3. FISHER, R. A. *Statistical methods for research workers*. Edinburgh: Oliver & Boyd, 1946.
4. HIGHBERGER, RUTH. The relationship between maternal behavior and the child's early adjustment to nursery school. *Child Developm.*, 1955, 26, 49-61.
5. KELLER, E. D. An exploratory study of nurturant behavior in young children. Unpublished master's thesis, State Univer. of Iowa, 1959.
6. McCANDLESS, B. R., CASTANEDA, A., & PALERMO, D. C. Anxiety in children and social status. *Child Developm.*, 1956, 27, 385-391.
7. McCANDLESS, B. R., & MARSHALL, HELEN R. A picture sociometric technique for preschool children and its relation to teacher judgments of friendship. *Child Developm.*, 1957, 28, 139-147.
8. MANN, NANCY. An investigation of dependency behavior in preschool children as related to maternal behavior and attitudes. Unpublished master's thesis, State Univer. of Iowa, 1959.
9. MILTON, G. A. A factor analysis study of child-rearing behaviors. *Child Developm.*, 1958, 29, 381-392.
10. MOORE, SHIRLEY G. Evaluation of a picture sociometric technique with young children. Unpublished manuscript, State Univer. of Iowa, 1958.
11. SCHAEFER, E. S., & BELL, R. Q. Development of a parental attitude research instrument. *Child Developm.*, 1958, 29, 339-361.
12. SEARS, R. R., MACCOBY, ELEANOR E., & LEVIN, H. *Patterns of child rearing*. Evanston: Row, Peterson, 1957.
13. THORPE, J. G. An investigation into some correlates of sociometric status within school classes. *Sociometry*, 1955, 18, 49-61.
14. ZUCKERMAN, M., RIBBACK, BEATRICE, MONASHKIN, I., & NORTON, J. A. Normative data and factor analysis on the parental attitude research instrument. *J. consult. Psychol.*, 1958, 22, 165-171.

CHILDREN'S ART ABILITIES: DEVELOPMENTAL TRENDS OF ART CHARACTERISTICS¹

BETTY LARK-HOROVITZ* and JAMES NORTON

Purdue University

The developmental nature of characteristics of children's drawings has been investigated extensively in the past literature on this subject. Other investigators have noticed, in a general manner—and in a number of case studies, especially of preschool children—that different characteristics show developmental trends of different nature. However, previous investigations have not attempted to observe the occurrence of a large number of art characteristics at each of a systematic sequence of age levels, as this study undertakes to do.

The data used here were collected during the years 1936 to 1940 at the Cleveland Museum of Art in experiments connected with the study of children's art abilities. One thousand and 15 children, 6 to 15 years old, of the Museum's Saturday classes from all parts of Greater Cleveland constitute the subjects employed in this investigation. There were approximately 50 boys and 50 girls at each age level, with only minor variations. From age 11 on, the subjects employed in this study belonged to two groups: gifted and nongifted children. The term "gifted" refers specifically to giftedness in regard to drawing, painting, or other creative activities in the field of visual arts. Because the range of "giftedness" within the "gifted" group was great, this group can be considered as a continuation of the "nongifted" children, the dividing point being to some extent arbitrary. The proportion of "gifted" children was as follows: age 11, 8 per cent; age 12, 4 per cent; age 13, 12 per cent; age 14, 20 per cent; age 15+, 33 per cent.

The children were asked to make pictures on seven subjects, the first four of which were analyzed. The analysis followed definitions set up for the various characteristics in order to minimize errors of judgment,² and described a number of single characteristics that could be perceived in terms such as Considerable Motion, Bandlike Arrangement, Hesitating Line, etc. Some of the data thus obtained are being used in the present paper which is concerned with the developmental trends of characteristics observed in the drawings mentioned above.

* Division of General Studies, School of Science, Purdue University, Lafayette, Indiana.

¹ Carried out under a grant of the General Education Board, New York, to the Department of Education of the Cleveland Museum of Art, Dr. Thomas Munro, Curator. The authors wish to express their appreciation to Dr. Munro and the Cleveland Museum of Art for permitting the use of these data.

² The recording of characteristics was carried out by Lark-Horovitz, Barnhart, Harding and Sills, under supervision of Lark-Horovitz.

CHILD DEVELOPMENT

This investigation proceeds in two directions (with results that permit comparisons): one is that of expressing developmental trends, and in terms more exact than in a previous publication (2); the other is that of expressing developmental trends in their relation to the single characteristic of Representation (see below), without reference to age or sex, but with reference to subject matter, a kind of study not carried out previously.

Description of Characteristics (Variables)

Developmental trends in the categories of each characteristic or sub-characteristic were studied by making use of one, two, three, or all four drawings, depending on the data available for a given characteristic. The subjects of these drawings were: I. Make a picture of whatever you like best to draw; II. Draw a picture of a man; III. Make a picture showing just how your classroom looks while the children and the teacher are there; IV. Make a picture showing what you would like to do next summer. The children were classified by sex and age so that the differential effects of these factors on development could be studied in the analyses reported below under the heading "Developmental Trends in Terms of Chronological Age." The same definitions of characteristics, their divisions and subdivisions, as described below, serve also for the analyses reported under the heading "Developmental Trends in Terms of Representational Stage," as far as these analyses make use of those characteristics.

The various characteristics under observation have been grouped, as seen below, according to some general qualities in drawing which seem to be basic to this grouping; and in discussing the results of this investigation we shall refer to them as A, B, C, D, and E.

A describes characteristics common to the general manner of Representation.

B deals with characteristics pertaining to Color qualities.

C treats of qualities inherent in Composition.

D contains characteristics related to the representation of Motion and therefore described as Motion.

E has to do with characteristics specifically involved in the treatment of the Medium that was used.

A. Representation

a. General Characteristics of Representation divided into Representational Stages

1. Preschematic Stage. This is a stage in children's drawings which follows the early scribble activities (characterized by lines or paint streaks distributed arbitrarily and in a noncoordinated manner over a surface without conveying meaning) and which combines characteristics of both, scribble and schema, but especially shows the step-wise development of schemata of various objects.
2. Schematic Stage. Drawings in this stage make use of a number of observations of essential characteristics of objects, combining often two or more points of vantage from which the object is observed, making use of whatever essentials are known and remembered, have roused the

BETTY LARK-HOROVITZ and JAMES NORTON

child's interest and which, at the moment of the drawing, may also have emotional overtones. Schemata lack correct proportions of the parts of a given object and also do not observe the "correct" proportions between given objects that constitute a picture. Side views, front views, and bird's-eye views are frequently combined.

3. Mixed Stage. This is a phase between 2 and 4, leading from the former to the latter, both in picturing of single objects and in the manner of arrangement of objects in a composition.
4. True-To-Appearance Stage. Representations of this kind give a view of an object or of several objects and of their compositional arrangement that is taken from a single point of vantage, show objects—"as they appear"—that is, only those parts that can be visible from a single point of vantage; proportions of the parts of an object and among various objects, when assembled and shown together, are fairly "correct."
5. Perspective Stage. In addition to the characteristics of 4, representation of space is added in form of distance perspective (line perspective) and foreshortening (i.e., the contracted appearance of certain dimensions of an object when the object does not lie in a plane perpendicular to the line of vision).

b. *Figure Characteristics*

1. Occurrence of the Human Figure as picture object
 - (a) Entire Figure
 - (b) Head (and Shoulders) only
2. Manner of picturing the Human Figure
 - (a) Position of Figure
 - (1) Front
 - (2) Profile
 - (3) Mixed Profile, that is a combination, in varying degrees, of front and side views of the same figure and face
 - (4) Three-Fourth View
 - (b) Sex Characteristics
 - (1) Masculine
 - (2) Feminine
 - (3) Childlike

B. *Use of Color*

a. *Multicoloredness*

1. Monochromes, that is drawings made in one color or in pencil
2. Two or Three Colors
3. Four or More Colors

b. *Local Coloring*

1. True Local Color, that is the actual color as part of a given object without differentiation of shadings or variety (i.e., water is blue, sky is blue: the same blue can be used for both; grass is green, leaves are green: the same green serves for both)
2. Variation of Local Color, that is two or more varieties in a given color (such as bluish green or yellowish green for varieties of leaves)

c. *Brightness of Color*, that can be described by the presence of a certain luster or luminosity of color effect

d. *Remarkable Use of Color*, marked by brilliancy or contrast, by integration of various colors, or by subtle blending, all of which differ from and surpass the usual application of color

CHILD DEVELOPMENT

C. *Composition*

a. *Grouping* (of three or more objects)

1. Nonorganized, that is showing no organization in their mutual arrangement within the picture
2. Bandlike Organization, that is an arrangement of objects in a row as on a line or flat strip, sometimes two or more strips, placed one above the other
3. Organized, that is arranged in a manner which indicates a balance or relation among the objects
4. Remarkably Well Organized

b. *Arrangement*

1. Symmetry or Near-Symmetry, that is an arrangement of the objects within the picture to the right and left of a perpendicular axis (distinctly imaginable), either side serving as a kind of mirror to the other or, in point and counterpoint balance, shifted on each side of the axis
2. Intentional Asymmetry, an arrangement which is nonsymmetrical, yet not haphazard; but organized though in an irregular manner

c. *Details* of objects which are meant to go beyond the basic necessities of picturing

d. *Surroundings*, that is indication of or extensive amount of objects in a composition beyond the main subject matter

e. *Shapes*

1. Clearly Defined Shapes, that is shapes that are distinctly contoured or outlined
2. Intentionally Indefinite Shapes, that is objects so represented that they appear blurred or as if fusing into another object without visible distinction; yet such indefiniteness is intended for a specific effect

f. *Style*

1. Naturalistic Style or resemblance to the "real" objects, therefore in the manner of a Naturalistic Style
2. Decorative Style, that is an ornamental treatment of shapes, contours, colors, grouping, etc., with repetitions that suggest a pattern
3. Expressionistic Style, insofar as it makes use of intentional distortions of shapes and colors, thus resembling certain kinds of "expressionistic" art

D. *Motion*

a. *Representation of Motion* in general

1. Motion Attempted and Indicated
2. Considerable Motion

b. *Execution of Motion*

1. Unsuccessful Attempts at representing Motion, whether slight or considerable
2. Successful representation of Motion, slight or considerable

c. *Flexibility*

1. Slightly flexible as contrasted to rigidity
2. Very flexible

E. *Use of Medium*

a. *Consistency*

1. Consistent Use of Medium, that is the use of pencil and color crayons in drawing and coloring in a manner which attempts the outline of

BETTY LARK-HOROVITZ and JAMES NORTON

shapes large or small enough to be consistent with the fine or blunt point of the medium and with the rough or smooth paper surface; and the shading and coloring of areas that befit the extent and limitations of these media

2. Consistent and Moderately Effective Use of Medium, that is an achievement of effects which goes beyond consistency alone, but is not extraordinary
3. Consistent and Remarkably Effective Use of Medium, that is an achievement of effects that is far beyond mere consistency of medium

b. *Area Treatment*

1. Ragged or Smudgy, that is a treatment of colored or penciled areas that give a ragged or smudged appearance, distinctively due to lack of skill or to carelessness
2. Smooth or Very Smooth, that is the application of color crayons (or pencil) in such a manner that the area seems smooth and finished
3. Differentiated, that is a treatment of areas that goes beyond smoothness, that is blended or graded, textured or bold; colors blend smoothly into each other or show gradation in application or a relevant texture such as controlled roughness to indicate some specific material; or a use of color which is daring, yet controlled such as an extreme but also brilliant red area

c. *Line Treatment*

1. Ragged, Hesitating, Weak
2. Decisive
3. Bold or Subtle, or both, beyond being decisive.

PROCEDURE

Each analysis involves, potentially, four factors: Sex, Age (or Representational Stage in the second set of analyses), Subject matter of Drawing (if more than one drawing is involved), and finally the categories recorded for the characteristic in question, which may represent *degrees*, as in the case of Consistency of Medium, or may represent *different manifestations* of the characteristic, as in the case of Style, or both, as in the case of Representation. To begin with, the children were classified as to Sex and Age (age groupings were by years from 6 to 14, and then one group for 15 or over). This defines 20 groups of children which we shall refer to as the Sex-Age groups. (In the second part these are replaced by the four groups referred to as Representational Stage groups.) Then for *each* drawing to be used in the analysis of the given characteristic, a tally was made of the number of children in each Sex-Age (or Representational Stage) group whose drawings on that Subject matter were classified into each of the several categories of that characteristic. This resulted in a frequency table giving the number in each Sex-Age (or Representational Stage) group whose drawings fell into each category of the characteristic for each Drawing (i.e., Subject matter). These frequencies were then changed into the *proportion* (or per cent) of each Sex-Age (or Representational Stage) group whose drawing

CHILD DEVELOPMENT

fell into each category of the characteristic, for each Subject matter. These proportions were then transformed by the angular transformation (3):

$$Y = \arcsin \sqrt{p}.$$

(The benefit from this transformation is that the sampling distribution of Y is more nearly normal than that of p , and this is desirable for the analysis of variance which is to be performed.)

For each characteristic an analysis of variance of the table of transformed proportions was carried out. Each such analysis comprised (potentially) sums of squares (and mean squares) for Sex, Age (or Representational Stage), Drawing (or Subject matter), if more than one was involved, Categories of the characteristic under analysis, and the interactions among those factors of various orders. In reporting the analyses in Tables 1a and 1c however, only the Category main effect and the interactions involving the Category factor are reported, since we are concerned here with the categories, and the influence of the other factors upon them. The Sex, Age, and Drawing main effects, and their interactions, are of no interest to us (and are in some cases meaningless anyway because of the way in which the proportions have been taken). The analyses reported in Table 1b involve analyses of Characteristics which have no subdivision into categories, and these are reported in full. In each case the highest order interaction mean square was used as the denominator in the F tests shown. Similar considerations apply to Table 2 reporting the analyses in terms of Representational Stages.

There are some recognizable shortcomings to this method of analysis: most important, from a theoretical point of view, is that the proportions are based on varying numbers of children in the different Sex-Age (and Representational Stage) groups, so that the sampling variances of the transformed proportions are not constant over the table. Also, the same children are represented by their drawings of different subject matters, so that a certain amount of nonindependence is present in the table. It is thought, however, that, despite these shortcomings, the analyses may be taken as indicative of meaningful differences in developmental trends and the effects on these trends of the various factors included.

RESULTS

The discussion and interpretation of results will be divided into two parts: the first will be concerned with the effects of Sex, Age and Subject matter; the second will deal with the effects of Representational Stages and Subject matter. In both parts the question is one of the effects of these factors upon the percentage of the drawings falling in the categories of the various characteristics. In the text we describe the nature of the more important effects found significant in the analyses of variance of Tables 1 and 2. A number of these are illustrated by accompanying figures. Others are described only in the text.

TABLE I
ANALYSES OF VARIANCE FOR DEVELOPMENTAL TRENDS
IN TERMS OF CHRONOLOGICAL AGE

(a) ANALYSES OF CHARACTERISTICS WHICH HAVE DIVISIONS INTO
CATEGORIES AND INVOLVE MORE THAN ONE DRAWING

Source	Represent. (Aa) 4 draw., 6-15+		Human Fig. (Ab1) 2 dr., 6-15+		Color (Ba) 4 dr., 6-15+	
	MS*	F*	MS	F	MS	F
Between Categories	35,552	1,575.90	23,978	357.23	38,400	8,402.51
Sexes × Categories	95	4.20			1,346	294.57
Ages × Categories	1,365	60.49	438	6.52	1,388	303.65
Drawings × Categories	310	13.73	9,353	139.34	942	206.21
S × A × C	57	2.50†			69	15.05
S × D × C					119	26.13
A × D × C					90	19.62
S × A × D × C	22.6		40.1		4.6	

Source	Color (Bb) 4 dr., 6-15+		Composition (Ca) 3 dr., 6-15+		Compos. (Cb) 4 dr., 6-15+	
	MS	F	MS	F	MS	F
Betw. Categories	105,781	4,167.90	2,592	115.20	5,954	91.72
S × C	616	24.29	108	4.84		
A × C			150	7.62	225	3.46
D × C	362	14.27	198	8.86	1,439	22.17
S × A × C	91	3.59				
S × D × C			51	2.27		
A × D × C	80	3.15	83	3.71		
S × A × D × C	25.4		22.4		64.9	

Source	Composition (Ce) 4 dr., 6-15+		Composition (Cf) 4 dr., 6-15+		Motion (Da) 4 dr., 6-15+	
	MS	F	MS	F	MS	F
Betw. Categories	40,513	1,347.00	9,980	212.15	20,070	770.93
S × C	1,040	346.70	231	4.69		
A × C	266	81.49	746	15.85	246	9.46
D × C	941	288.76	884	18.80	163	6.26
S × A × C						
S × D × C					141	5.46
A × D × C						
S × A × D × C	3.3		47.0		26.0	

Source	Motion (Db) 4 dr., 6-15+		Medium (Ea) 4 dr., 6-15+		Medium (Eb) 4 dr., 6-15+		Medium (Ec) 4 dr., 6-15+	
	MS	F	MS	F	MS	F	MS	F
Betw. Categ.	21,060	1,316.25	28,388	362.18	9,340	314.06	12,871	707.60
S × C								
A × C	56	3.55‡	630	30.23	677	22.77	412	22.68
D × C	470	29.38	326	15.65	620	20.85		
S × A × C							70	3.82
S × D × C			94	4.50				
A × D × C	232	14.47	57	2.71				
S × A × D × C	15.9		20.8		29.7		18.2	

* Mean squares and *F* ratios are given only for effects significant at the .01 level.

† Italicizing of an MS and *F* (such as 57 and 2.50 in Aa) indicates illustrative figure is provided.

‡ This particular MS and *F* is given for an effect significant at the .05 level.

CHILD DEVELOPMENT

TABLE 1 (continued)

(b) ANALYSES OF CHARACTERISTICS WHICH HAVE NO DIVISIONS INTO CATEGORIES

Source	Color (Bc) 4 dr., 9-15+		Color (Bd) 4 dr., 6-15+		Compos. (Cc) 2 dr., 6-15+		Compos. (Cd) 3 dr., 6-15+	
	MS	F	MS	F	MS	F	MS	F
Betw. Sexes	141	43.52			429	101.42		
Betw. Ages	90	9.61	80	3.46	60	14.21		
Betw. Drawings ..	3,141	335.89					4,406	6.38
S × A					17	3.97		
S × D					189	44.80		
A × D	126	13.49			56	13.14	147	2.14
S × A × D	9.4		23.3		4.2		69.2	

(c) ANALYSES INVOLVING ONE DRAWING ONLY

Source	Represent. (Ab2(a)) 6-15+		Represent. (Ab2(b)) 6-15+		Motion (Dc) 6-15+	
	MS	F	MS	F	MS	F
Betw. Categories	5,698	67.67	5,253	52.01	3,478	131.40
Sexes × Categories ..	140	16.60	1,590	15.74		
Ages × Categories	131	15.55	346	3.43	99	3.73
S × A × C	8.4		100.7		26.5	

Developmental Trends in Terms of Chronological Age

Generally speaking, for the characteristics investigated here, there are strong trends associated with the age factor, i.e., developmental trends. Many of the characteristics also show significant sex effects and interactions, but in most cases these are not strong enough to obscure the general trend associated with the age factor even where there is a significant sex-age interaction. The four different subject matters or topics that constitute the content of the drawings also usually show significant effects but, as with the sex effects, these subject matter effects usually do not interact with age strongly enough to produce important deviations from the general age trends of the characteristics under investigation.

The age trends of the five Stages of Representation, Aa, are comparatively simple (Fig. 1): Stage 1 decreases rapidly from age 6 to 8 after which it occurs only sporadically; Stage 2 takes an upward surge to age 7 (for girls) or 8 (for boys), and then decreases almost steadily until age 15+.³ Here, however, one can observe clearly an example of the kind

³ The sign + designates a limited number of children over 15 years old who were included in the 15-year-old group.

BETTY LARK-HOROVITZ and JAMES NORTON

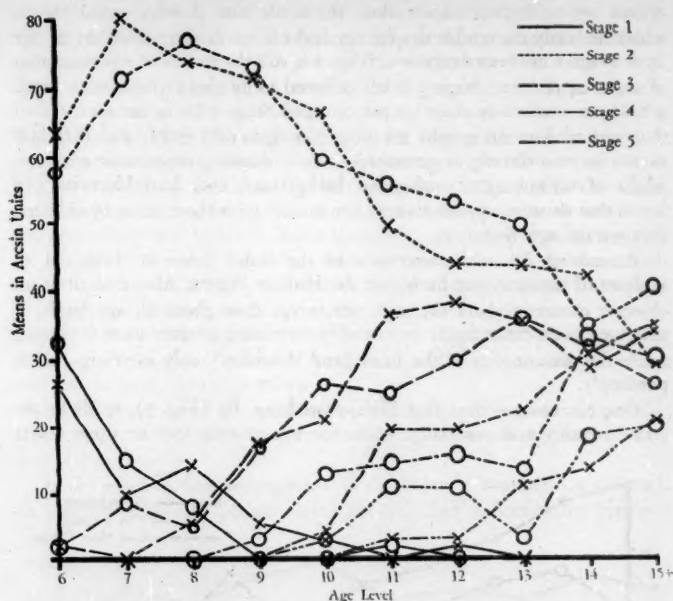


FIGURE 1—Developmental trends in terms of chronological age: Representational Stages 1, 2, 3, 4, and 5 (Aa); S \times A \times C. (Note: Data for boys are indicated by "o"; data for girls, by "x.")

of sex variation which does not alter the general trend, but rather demonstrates a shift of achievement in the sexes at two or more ages: the boys reach a peak at 8, the girls at 7; boys show a more definite drop at age 14, girls maintain a plateau between 11 and 14. Representational Stage 3 has a rising curve which reaches its peak at 12 to 13 and then decreases; it also has a distinct "sex-shift" of about one year between the peak for girls and the peak for boys. Stage 4 starts at ages 9 and 10, maintains a plateau between 11 and 13, then follows again an upward trend; here boys develop at first faster than girls, then girls faster than boys, yet the general trend remains clear. Stage 5, reached by a comparatively small number of children, shows a slow but steady increase. (It should be kept in mind that this experimental result may be due to the increasingly larger percentage of "gifted" children of these age levels who attend Saturday classes at the Museum.)

While Stage 1 of Representation vanishes with increasing age, Stage 2 develops a "high" early and decreases from there on; Stage 3 increases up to later age levels and then decreases; Stages 4 and 5 are still on the increase

at age 15+. Representation thus shows definite developmental trends which are early discernible despite sex and subject matter variation: at age 15+ Stage 1 has been outgrown, Stage 2 is still the means of representation of about 24 per cent, Stage 3 is still adhered to by about 31 per cent, Stage 4 has been reached by about 31 per cent and Stage 5 by 14 per cent. (Note that vertical axes on graphs are in arcsin transform units, and therefore cannot be read directly as percentages.) In a drawing experiment with 170 adults of various ages, professions, background, etc., Lark-Horovitz (1) found that drawings by adults compare mainly with those made by children between the ages 6 and 10.

Examining other characteristics in the order listed in Table 1a of analyses of variance, one finds that the Human Figure, Ab1, though as an object it occurs in about the same percentage throughout all age levels, is presented as an entire figure in a steadily decreasing proportion of drawings, while the presentation of the head (and shoulders) only increases correspondingly.

One can observe next that Multicoloredness, Ba (Fig. 2), tends to decrease steadily and markedly, while the use of only two or three colors

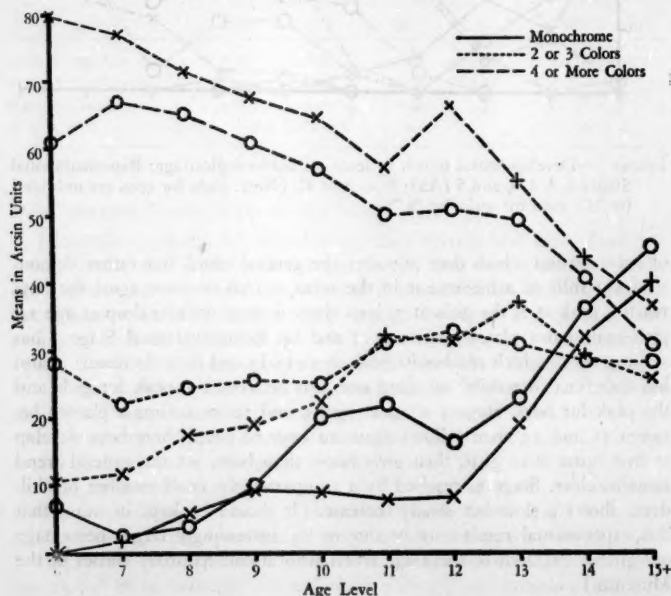


FIGURE 2—Developmental trends in terms of chronological age: Multicoloredness (Ba); S \times A \times C. (Note: Data for boys are indicated by "o"; data for girls, by "x.")

remains approximately stable; and that monochrome picturing has an upward trend, first gradual and slow, then with a sharp turn upwards between ages 13 and 15+. As has been pointed out when discussing the trends of Representational Stages, the sex variation here again shows shifts of achievement in the sexes without altering the general age trend. For example, in Use of Four or More Colors girls at age 9 are comparable to boys at age 7, girls at age 10 to boys at age 8, and girls at age 11 to boys at age 10. A further variation in the trends can be observed—associated with the different Drawings, but again without interference with the general trend, i.e., the general upward trend in Monochromes is quite steep for Drawing II, far less so for Drawing IV. When drawing essentially a single object such as in the Picture of a Man, children are inclined to concentrate on outlines of shapes and detailed attributes of the object which, in general, they are used to handle with a pencil; while a picture of summer activities, as in Drawing IV, suggests "pictorial" parts such as landscape, various objects, and people remembered in colors.

In the Use of Local Colors, Bb (Fig. 3), though the age trend is significantly different for the two sexes, no significant over-all age trends (i.e., $A \times C$) appears.

In the compositional arrangement of Organized Grouping, Ca (Fig. 4), an upward trend can be seen until age 11, then a fluctuation within a

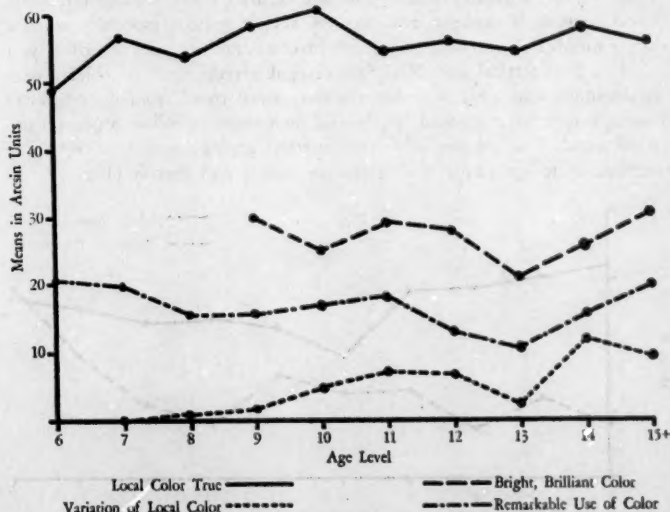


FIGURE 3—Developmental trends in terms of chronological age: Use of Color (Bb, Bc, Bd); $A \times C$. (Note: Bright, Brilliant Color not recorded before age of 9.)

CHILD DEVELOPMENT

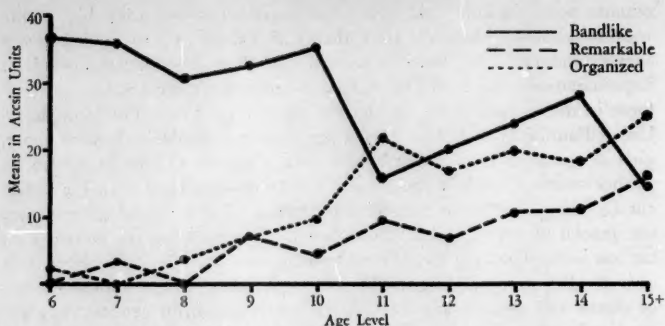


FIGURE 4—Developmental trends in terms of chronological age: Grouping (Ca); $A \times C$.

plateau for the remaining four age levels; in Remarkably Well Organized Grouping, on the other hand, there is a later start (age 9), a lower percentage of children are involved, and the curve oscillates within a narrow range increasing at age 15+. The curve for Bandlike Grouping, Ca₂, remains on a plateau at 30 per cent between ages 6 and 10, drops sharply at age 11, then increases steadily until age 14, after which it drops off. Bandlike Grouping is strongly influenced by subject matter, especially when a larger number of people or objects are involved, and it varies accordingly.

The Symmetrical and Near-Symmetrical arrangement of objects in a composition, Cb₁ (Fig. 5), shows a downward trend from ages 6 to 10—at age 10 it has a marked “dip”—and then seems to follow a general upward trend. The Intentionally Asymmetrical arrangement, Cb₂, seems to oscillate up to age 12 and then increases steadily and sharply (Fig. 5).

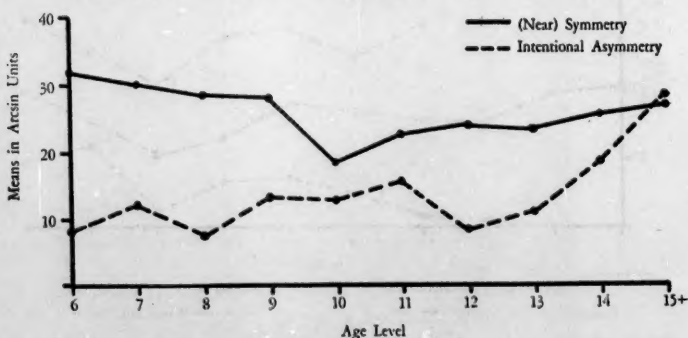


FIGURE 5—Developmental trends in terms of chronological age: Arrangement in Composition—Symmetry and Asymmetry (Cb); $A \times C$.

Clarity of Shapes, Ce, tends to remain on a plateau, though it oscillates slightly between ages 7 and 13, then increases sharply, while the Intentional Indefiniteness of Shapes shows a tendency to increase, though only to a slight extent.

The various categories of resemblance to Style, Cf, have significantly different age trends: drawings that show a resemblance to a Naturalistic-Realistic Style increase continuously after age 7, sharply at age 14; drawings that bear resemblance to an Expressionistic Style are relatively few and seem to remain on the same level until age 14, then increase definitely at age 15+; drawings that give an impression of a Decorative Style have a decreasing trend through all age levels.

The curve for Attempted Motion, Da1, remains relatively flat, with only slight oscillation. The curve for Considerable Motion, Da2, shows a slightly increasing trend (regardless of the success in representing it).

In Fig. 6 one sees that Successfully Represented Motion, Db, shows a slight upward trend with a sharper increase at age 15+.

The technique of drawing as expressed by Consistent Use of Medium, Ea (Fig. 7), also shows age trends: Consistent Use of Medium, Ea1, remains about the same—with slight aberrations—until age 13, but then drops at ages 14 and 15; Consistent Use of Medium but with Moderate Effectiveness, Ea2, has an upward trend, with a particular high at age 14; and Remarkably Effective Use of Medium, Ea3, is confined to a relatively small percentage of children and shows an upward trend at ages 14 and 15+.

In regard to Treatment of Areas, Eb, the category Ragged Area Treatment shows a steadily decreasing trend; Smooth Area Treatment a slight

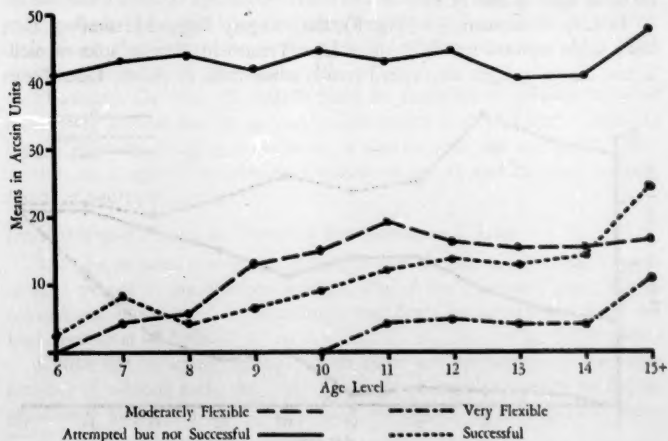


FIGURE 6—Developmental trends in terms of chronological age: Motion—Execution (Db)—Flexibility (Dc); A \times C.

CHILD DEVELOPMENT

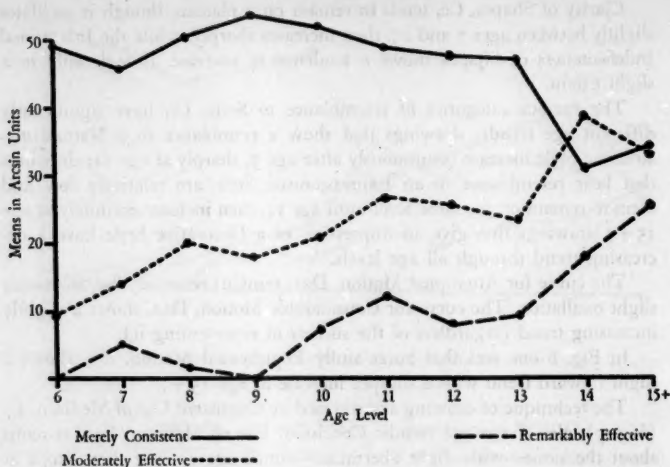


FIGURE 7—Developmental trends in terms of chronological age: Consistency in Use of Medium (Ea); A x C.

upward trend to age 11 and then remains on a plateau; while Differentiated Area Treatment, represented by a relatively small number of children, increases at ages 9 and 10, remains then on a plateau, and shows a small increase at ages 14 and 15+.

In Line Treatment, Ec (Fig. 8), the category Ragged-Hesitating, Ec1, has a slight upward trend; Decisive Line Treatment, Ec2, in spite of oscillation, shows a slight downward trend, while Bold or Subtle Line Treat-

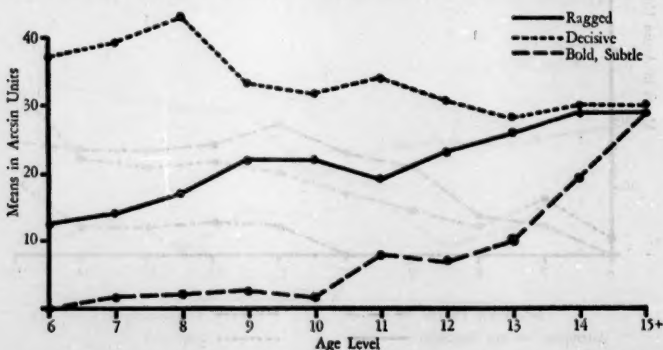


FIGURE 8—Developmental trends in terms of chronological age: Line Treatment (Ec); A x C.

ment, Ec3, remains on a plateau for the first five age levels but has a sharp upward trend at ages 14 and 15+.

Table 1b gives the analyses of characteristics which have no division into categories. These will be discussed in the order given in the table. Use of Bright Color, Bc (Fig. 3), shows no definite age trend, though the age effect is significant, apparently due to oscillation; Use of Bright Color, however, varies according to subject matter, especially in Drawing I where it has a definite downward trend. This is also the case in Remarkable Use of Color, Bd (Fig. 3), which fluctuates within a limited range of 12 per cent and 3 per cent, and is confined to a small number of children.

Details in Composition, Cc, besides showing a significant difference between sexes, also has a significant age effect which is due to oscillation rather than any clear developmental trend. Surroundings, Cd, do not show any over-all age trends.

Table 1c lists analyses involving one drawing only. The Position in which the Human Figure, Ab2(a), is represented shows well defined trends: there is a downward trend of the front view to age 10 coinciding with an upward trend of profile views; while the profile views decrease from then on, the front views increase; Mixed Profile which occurs increasingly from ages 6 to 8 and then decreases, and the Three-Fourth view, which has an upward trend from age 9 on, are both limited to a smaller number of children. Sex Characterization, Ab2(b), which is a representation of masculine or feminine (or child) characteristics, follows a developmental pattern. Masculine characterization has an over-all upward trend, Feminine characterization ascends until age 13, then falls off, while Child characterization has a peak at age 9, then decreases steadily. There is a great difference in these characterizations between boys and girls: nonexistent interest in feminine characterization on the part of the boys, great interest in child characterization on the part of the girls.

Flexibility, Dc (Fig. 6), which could be described as a latent form of motion, is represented by a very small number of children: Moderate Flexibility starts at age 7, remains on a plateau after age 11; Great Flexibility starts at age 11, remains on a plateau to age 14 and increases to some extent at age 15+.

Developmental Trends in Terms of Representational Stage

The characteristics investigated in this part do not cover all the characteristics treated in the previous sections. Use of the Human Figure, Multicoloredness, Details and Surroundings, and Sex Characterization have not been treated at all here.

Unlike the 20 sex-age groups which contained approximately the same number of subjects each, the 3,982 drawings arranged according to Representational Stage (1+2),⁴ 3, 4, 5, have an entirely different distribution:

⁴ Representational Stage 1 and 2 which were kept separate in the study of developmental trends in terms of chronological age have been combined here and, in order to avoid confusion, have been described as Stage (1+2).

CHILD DEVELOPMENT

TABLE 2

ANALYSES OF VARIANCE FOR DEVELOPMENTAL TRENDS
IN TERMS OF REPRESENTATIONAL STAGE

Source	Color (Bb, c, d) 4 dr., 6-15+		Compos. (Ca) 3 dr., 6-15+		Compos. (Cb) 4 dr., 6-15+		Compos. (Ce) 4 dr., 6-15+	
	MS	F	MS	F	MS	F	MS	F
Betw. Categories	3,171	65.52	127	3.18				
Representational Stages \times Cat.	482	9.96	298	7.45	528	16.58		
Drawings \times Cat.	130	2.69			407	12.77	457	3.11
R \times D \times C	48.3		40.0		31.9		146.9	

Source	Compos. (Cf) 4 dr., 6-15+		Motion (Db) 4 dr., 6-15+		Medium (Ea) 4 dr., 6-15+		Medium (Eb) 4 dr., 6-15+		Medium (Ec) 4 dr., 6-15+	
	MS	F	MS	F	MS	F	MS	F	MS	F
Betw. Cat. ...	4,155	301.54			765	30.31	266	9.99	1,426	23.85
R \times C	1,182	85.76	1,145	9.08	1,577	62.54	852	32.07	1,269	21.22
D \times C	495	35.93					121	4.56		
R \times D \times C	13.8		126.1		25.2		26.6		59.8	

Representational Stage (1+2) covers 69 per cent of all drawings, Stage 3 covers 18 per cent, Stage 4 covers 9 per cent, and Stage 5 covers only 4 per cent of all drawings. Thus, percentages occurring in Stage 5 especially must be viewed with this distribution in mind.

Following the same order of discussion as in Part I, i.e., starting with Table 2 of analyses of variance, and referring to the sequence of Representational Stages as: (1+2), 3, 4, 5, one can observe that Use of True Local Color, Bb1, has a decided downward trend; Variation of Local Color, Bb2, occurs only in Stage 3 and—increasingly—in Stage 4; Bright and Brilliant Color, Bc, starts already at Stage (1+2), increases in each of the subsequent stages slightly, but decreases in Stage 5; and Remarkable Color Use, Bd, has an upward trend through all four stages (Fig. 9).

In regard to Grouping, Ca (Fig. 10), one can follow a continuous down-trend of the Bandlike Grouping, an increase of Organized Grouping from Stage (1+2) to 3, then a plateau-like level to Stage 5. Remarkably Well Organized Grouping increases steadily through all four stages.

The trend of Symmetrical Composition, Cb (Fig. 11), is slightly downward, that of Intentional Asymmetry, steeply upward.

Clearly Defined Shapes and Intentionally Indefinite Shapes, Ce, both have an upward trend accounting for nonsignificant R \times C interaction.

Concerning Style, Cf, the proportion of drawings bearing resemblance to a Naturalistic-Realistic Style increases suddenly from Stage (1+2) to Stage 3, still increases decidedly to Stage 4 and then falls off slightly; the proportion of drawings that suggest an Expressionistic manner have a steady upward trend, while Decorative drawings have a downward trend.

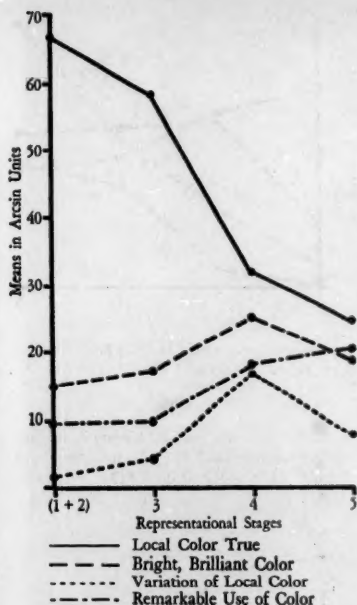


FIGURE 9—Developmental trends in terms of representational stage: Use of Color (Bb, Bc, Bd); R×C. (Cf. Fig. 3.)

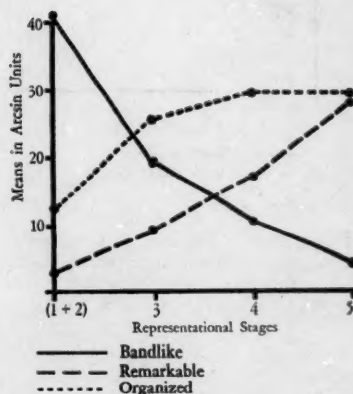


FIGURE 10—Developmental trends in terms of representational stage: Grouping (Ca); R×C. (Cf. Fig. 4.)

Unsuccessful Motion, Db (Fig. 12), decreases slowly and continuously; Successful Motion shows a steady, sharply increasing trend.

In technical ability, the Consistent Use of Medium, Ea (Fig. 13), which does not exceed plain consistency, is high at Stage (1 + 2)—62 per cent—and drops off steeply and steadily to 9 per cent. Consistent Use of Medium with Moderate Effectiveness has an ascending trend up to Stage 4 and then falls off slightly; while Remarkable Effectiveness in the use of medium increases from Stage (1 + 2), and very sharply from Stages 4 to 5.

Ragged Treatment of Areas, Eb, has a steep downward trend up to Stage 4, then slopes more gently; Smooth or Very Smooth Area Treatment remains approximately the same through all stages; and Differentiated Area Treatment remains on a steep upward trend to Stage 4, then tapers off.

Of the three kinds of Line Treatment, Ec (Fig. 14), the Hesitant and Ragged Line which is about the same at Stages (1 + 2) and 3 plunges steeply down to Stage 5; Decisive Line remains approximately alike through all stages; while the use of Bold and Subtle Line increases first slightly, then, at Stages 4 and 5, sharply.

CHILD DEVELOPMENT

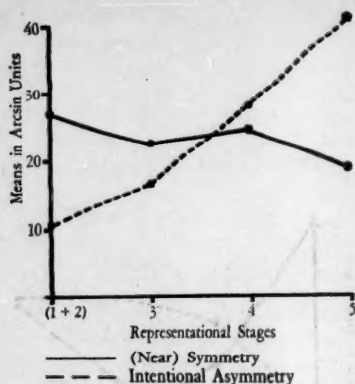


FIGURE 11—Developmental trends in terms of representational stage: Arrangement in Composition—Symmetry and Asymmetry (Cb); R x C. (Cf. Fig. 5.)

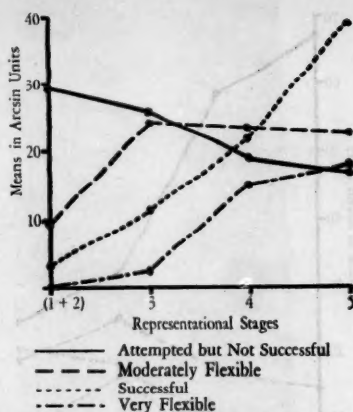


FIGURE 12—Developmental trends in terms of representational stage: Motion—Execution (Db)—Flexibility (Dc); R x C. (Cf. Fig. 6.)

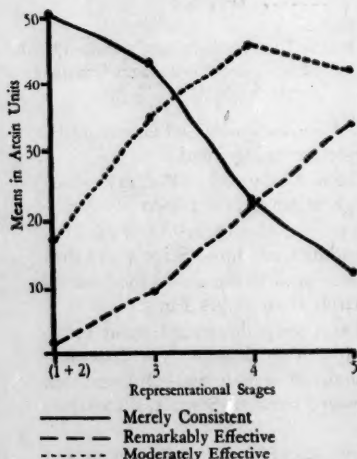


FIGURE 13—Developmental trends in terms of representational stage: Consistency in Use of Medium (Ea); R x C. (Cf. Fig. 7.)

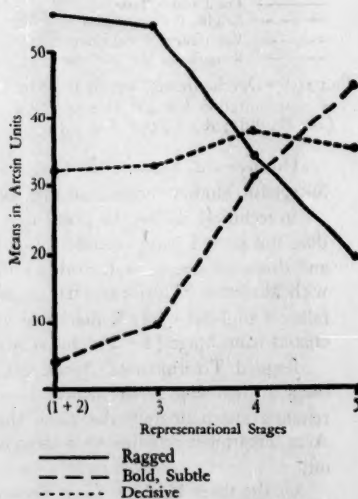


FIGURE 14—Developmental trends in terms of representational stage: Line Treatment (Ec); R x C. (Cf. Fig. 8.)

For the characteristics Position of Human Figure and Flexibility only one drawing was available for analysis. As a result, the Representational Stage \times Category interaction, which is the one of interest here, is the highest order interaction occurring, there being no interaction of higher order available to use as an error term in testing $R \times C$. For this reason, no analysis of variance table is presented for these characteristics. Nevertheless, the trend represented by $R \times C$ for the characteristic Flexibility, *De*, has been graphed for the purpose of comparison (Fig. 12). Flexibility, which has an upward trend to age 11 (Fig. 6), shows an upward trend both for Moderate and Strong degrees of Flexibility to Stages 3 and 4, then levels off. In regard to Position of the Human Figure, the Front View, *Ab2(a)*, shows a downward trend to Stage 5; the Profile View remains on approximately the same—high percentage—level; while Three-Fourth View ascends sharply through all stages; Mixed Profile occurs only during Stage (1 + 2).

The results of the study of developmental trends in terms of chronological age lead to the conclusion that, for most of the characteristics under investigation, the Age \times Category interactions are significant; that a large number of those interactions show distinct developmental trends, some of them throughout all 10 age levels here available, some through a lesser number of age levels. Even when significant Sex \times Age \times Category interactions occur, these developmental trends are not obscured by the sex difference. The most definite age trends are those of the Representational Stages which, by description and definition, are not illustrative of a single characteristic but rather of a conglomerate of characteristics. A few other characteristics show developmental trends that seem to parallel those of Representational Stages, such as Multicoloredness, which parallels the trends of Representational Stages 1 and 2, and Monochrome, which parallels Stages 4 and 5. Similarly, Use of Medium shows certain trends parallel to those of Representational Stages: plain Consistency of Medium, Ragged Area Treatment seem to be parallel to Representational Stages 1 and 2, while Remarkable Treatment in all three of these Uses of Medium shows age trends parallel to those of Stages 4 and 5.

Various Uses of Color, *Bb*, *Bc*, *Bd* (see Figs. 3, 9), Compositional characteristics such as Grouping, *Ca* (see Figs. 4, 10) or Shapes, *Ce*, show clearer developmental trends when plotted against Representational Stages than against Age levels. So also do characteristics such as Style, *Df*, Motion, *Db*, and the "remarkable" manifestations and degrees of a number of characteristics.

Although curves for the 10 age levels and for the four representational stages are not directly comparable, the nature of the various curves permits certain comparisons of trends.

The age trends of some of the art characteristics or their subcategories are quite explicit. However, the trends of a larger number of characteristics are more closely related to the sequence of Representational Stages.

Thus it seems that the occurrence and development of these art characteristics, in their different degrees and manifestations, is more closely related

CHILD DEVELOPMENT

to the sequence of Representational Stages than to chronological age, and may be more reliably predicted by using them.

SUMMARY

From data collected in connection with a diversified study of children's art abilities, certain characteristics of children's drawings (mainly concerned with representation in general, color qualities, compositional qualities, and qualities linked with the medium employed) were used for investigating their developmental nature.

The characteristics in question and their subcategories were studied for their interaction with Age, Sex and Subject matter and for their interaction with Representational Stages in drawing and Subject matter.

Analyses of variance were carried out and resulted in a number of significant F ratios which in turn served as indicators for the study of trends of characteristics. Graphing of the curves of categories of the various characteristics served as a basis for the observation of developmental trends of these characteristics in relation both to Age and Sex and the possible effect of Subject matter and in relation to Representational Stages and the influence of Subject matter.

The interpretation of the results leads to the conclusion that most interactions of Age and categories of Characteristics are significant and that a great number of those interactions show distinct developmental trends, in spite of frequent additional and significant interaction of Sex with Age and Characteristics. The significant interactions of Representational Stages with—mostly—the same Characteristics and subcategories that were used in the Sex-Age interaction study, that is characteristics concerning color, composition, motion and treatment of medium, point to clearer developmental trends than do the interactions with Age, especially in the "remarkable" manifestations of the various characteristics. Consequently, prediction of ability on the basis of the relation of characteristics to the sequence of Representational Stages seems more promising than that based on chronological trends.

REFERENCES

1. LARK-HOROVITZ, B. Interlinkage of sensory memories in relation to training in drawing. *J. genet. Psychol.*, 1936, 49, 69-89.
2. MUNRO, T., LARK-HOROVITZ, B., & BARNHART, E. N. Children's art abilities: Studies at the Cleveland Museum of Art. *J. exp. Educ.*, 1942, 11, 97-155.
3. SNEDECOR, G. W. *Statistical methods*. (4th Ed.) Ames, Iowa: Iowa State College Press, 1946.

SEX DIFFERENCES IN THE LIFE PROBLEMS AND INTERESTS OF ADOLESCENTS, 1935 AND 1957¹

DALE B. HARRIS*

The Pennsylvania State University

This is a study of interests and problems which uses the method of rank order rather than a check list as its research technique. The items ranked are 15 topics selected from the concerns of adolescents. High school students were asked first to consider the items as personal problems and to construct an order reflecting their own experience with the issues as personal problems. The students then reranked the same items in order of interest, considering the topics as things they would like to read about and discuss or hear discussed.

The problems and instructions for ranking were taken verbatim from a study published by Symonds (1, 2). He had selected the issues from young people's own discussions and phrased the issues in terms used by young people themselves. His 1641 students attended junior and senior high schools in Tulsa, Oklahoma, and New York City. The 1165 youth in the present study came from the junior-senior high school in a Minnesota community. Twenty-two years and considerable social, cultural, and economic change separate the circumstances of the two studies. Whether geographic or regional differences also influence the data cannot be known. In both studies, the samples represented in general the socioeconomic distribution of the communities from which they were drawn. The comparability of samples drawn from different geographic areas without close control through a stratification procedure is questionable. But the comparability of samples is also problematical across long time intervals, in which social and cultural changes have occurred, even when stratification by some socioeconomic index has been attempted. Changing conditions may themselves affect the index.

The hypothesis of the present study is taken from Symonds' discussion (2): "Change the social and economic structure of society and you immediately change the relative emphasis of these problems and interests" (p. 752).

In a ranking process as in any *system*, change in one feature or aspect may have widespread effects throughout the system. Thus, there are some

* Department of Psychology, Pennsylvania State University, University Park.

¹ This study was supported by the Institute of Child Development, University of Minnesota.

CHILD DEVELOPMENT

limitations to the method. The elements are *relative* to one another. The resulting rank order is not a true *scale* of values. Each choice made removes a degree of freedom and relates to the choices already made and those yet to be made. The ranks accorded a series of stimuli are systematically affected by the order in which the stimuli are first presented. Symonds removed this effect in his study by presenting the items in reverse order to approximately half his subjects. This study used the same procedure. The sample of children was drawn randomly from the available supply to constitute groups of 100 boys and 100 girls at each grade; half of each sex responded to the items in the order Symonds presented them in his report, and the other half responded to a reversed order.

For 15 items ranked in one experience, a rho value of .535 is required to establish a relationship greater than chance at the 5 per cent level of significance. To interpret differences in rankings accorded items by boys and girls, or changes between two sets of rankings made at different times, one can place more confidence in *mean* ranks of samples than in ranks made by individuals. But the meaning of the magnitude of the differences thus described is still elusive. By building a body of experience with different arrangements of the data, one can possibly increase the insight he brings to bear on any one arrangement.

For example, when means of rankings of items considered as problems, established by 100 boys in each grade, 7 through 12, are compared, the average or typical between grade-group rank order correlation is $+.74$ (estimated from Kendall's *W*). For the same items considered as interests, the typical between grade-group rank order correlation is $+.78$. For girls, the corresponding values are $+.70$ and $+.72$. Thus, children's rankings from grade to grade change no more on the average than is suggested by correlation values in the .70's. Putting the data of boys and girls together in each grade, thus increasing reliability through the size of *N* but also introducing any existing systematic sex differences, the values become $+.65$ (problems) and $+.63$ (interests), (both estimated from Kendall's *W*).

Because there are likely to be intrinsic relationships in any set of values ranked by a group of judges, simply because judges from similar backgrounds are not reacting blindly or randomly to meaningful material, it is instructive to examine some typical interarray correlations estimated from Kendall's *W*. If we consider the four ranks of mean values accorded by boys and girls in 1935 and 1957 to *problems*, this typical interarray correlation is $+.58$. For four arrays of *interests* this value is also $+.58$. Throwing together the eight arrays of problems and interests, the value becomes $+.36$. These values provide a pragmatic bench-mark to judge the meaning of the correlation values presented below, which have a theoretical top value of $+1.00$ and describe a chance relationship as $.00$.

Taking all boys regardless of grade and similarly combining all girls for purposes of comparison with data obtained on the same items by Symonds in 1935, some interesting observations can be made. Youth showed

DALE B. HARRIS

greater consistency in the ordering of these items, considered first as problems and then as interests, in 1935 than they do at present. This change is true both for boys (+.63 compared with -.23) and for girls (+.66 compared with +.25). Thus, in 1957 there is little correspondence between a set of issues considered as problems and the order of interest in the same issues. This was not true in 1935, when there was a noticeable correspondence between the sets of ranks,

But how do young people separated by almost a generation compare in the ordering of their problems and their interests? Boys' rankings of interests across the years are somewhat more consistent (+.76) than their rankings of problems then and now (+.47). For girls, there is little difference between relative positions accorded problems and interests over the years (problems, +.50; interests, +.55).

Do boys and girls accord the same order of importance to a set of adolescent problems? In both periods boys and girls rank their problems similarly. In 1935 the similarity of the rank order of the sexes is expressed by a correlation value of +.80. In the 1957 study the value is +.77. The similarity of boys and girls is as great, on the average, as the similarity of successive grade groups of boys, or of girls. When the items were ranked according to interest in 1935, a value of +.80 expresses the similarity of boys' and girls' judgments. The comparable figure in 1957 is +.58.

TABLE I

RANKS ACCORDED ISSUES CONSIDERED AS PROBLEMS BY HIGH SCHOOL BOYS AND GIRLS IN 1935 AND IN 1957

Issue	B O Y S				G I R L S			
	1935		1957		1935		1957	
	Mean Rank	Rank	Mean Rank	Rank	Mean Rank	Rank	Mean Rank	Rank
Health	6.7	(2)	9.1*	(12)	6.6	(2)	8.7*	(12.5)
Love, marriage	10.8	(15)	9.2*	(13.5)	11.0	(15)	8.5*	(10)
Safety	8.3	(8.5)	9.2*	(13.5)	8.8	(12)	10.0*	(14.5)
Money	6.2	(1)	6.3	(2)	6.8	(3)	6.5	(2)
Mental hygiene	8.7	(13)	8.2	(8)	8.2	(9.5)	6.9*	(3.5)
Study habits	6.8	(3)	5.0	(1)	7.4	(6)	6.3*	(1)
Recreation	8.3	(8.5)	10.1	(15)	8.4	(11)	10.0*	(14.5)
Personal, moral qualities	7.1	(4)	6.6*	(3)	7.3	(4)	7.1	(5)
Home, family relationships	8.2	(7)	8.4	(10)	8.2	(9.5)	7.6	(6)
Manners	8.5	(11.5)	7.6*	(5)	7.4	(6)	8.6*	(11)
Personal attractiveness	7.9	(6)	7.7	(6)	6.2	(1)	6.9*	(3.5)
Daily schedule	9.0	(14)	8.3*	(9)	9.4	(14)	8.7*	(12.5)
Civic interest	8.5	(11.5)	7.9*	(7)	9.0	(13)	8.4	(9)
Getting along with other people	8.4	(10)	8.6	(11)	7.9	(8)	8.0	(8)
Philosophy of life	7.5	(5)	7.5	(4)	7.4	(6)	7.7	(7)

* Change from 1935 significant at the 1 per cent level.

CHILD DEVELOPMENT

TABLE 2

RANKS ACCORDED ISSUES CONSIDERED AS INTERESTS BY HIGH SCHOOL BOYS AND GIRLS IN 1935 AND IN 1957

Issue	B O Y S				G I R L S			
	1935		1957		1935		1957	
	Mean Rank	Rank	Mean Rank	Rank	Mean Rank	Rank	Mean Rank	Rank
Health	5.6	(2)	6.4*	(2)	6.6	(4)	7.0	(6)
Love, marriage	9.3	(13)	7.7*	(8)	9.4	(12)	6.5*	(4.5)
Safety	7.8	(7)	8.7*	(11)	9.2	(10)	10.2*	(14)
Money	7.1	(3)	6.5*	(3)	8.1	(8)	8.2	(10)
Mental hygiene	9.6	(14)	9.5	(13)	9.8	(13.5)	8.0*	(9)
Study habits	8.7	(11)	9.4*	(12)	9.3	(11)	9.9*	(12.5)
Recreation	4.9	(1)	5.6*	(1)	5.6	(2)	7.8*	(8)
Personal, moral qualities	7.7	(5.5)	7.3*	(6)	7.6	(7)	7.1	(7)
Home, family relationships	8.4	(10)	7.2*	(5)	8.3	(9)	6.4*	(2.5)
Manners	7.5	(4)	8.6*	(10)	6.3	(3)	8.5*	(11)
Personal attractiveness	8.1	(8)	8.0	(9)	5.4	(1)	6.0*	(1)
Daily schedule	10.5	(15)	10.9	(15)	10.4	(15)	11.4*	(15)
Civic interest	9.0	(12)	9.6*	(14)	9.8	(13.5)	9.9	(12.5)
Getting along with other people	8.2	(9)	7.1*	(4)	7.0	(5)	6.4*	(2.5)
Philosophy of life	7.7	(5.5)	7.4	(7)	7.3	(6)	6.5*	(4.5)

* Change from 1935 significant at the 1 per cent level.

The changes described above become more interesting when we look at specific ranks in Tables 1 and 2. As problems (Table 1), three items change five ranks or more across time for both boys and girls (health, mental health, manners), three more for boys only (safety, recreation, schedule), and two more for girls only (love and marriage, and study habits). As interests (Table 2), three items change for both sexes across time (love and marriage, family relations, manners), and one more for each sex singly (recreation for girls, getting along with others for boys). There are as many sex differences now as in 1935. At that time seven topics significantly differentiated the sexes as problems at the .01 level of certainty.² Now eight topics satisfy the .01 level.

Sex differences in interests also are about the same. In 1935 boys and girls rated nine topics quite different. In 1957, 10 topics satisfied the .01 level.

Attention to item placements show that a number of topics are relatively high as sources of problems: *Money* is still high as a problem to

² This index of statistical significance refers to the difference between mean ranks, interpreted in terms of the standard errors of these means. It does not refer to shift in rank order.

DALE B. HARRIS

both boys and girls (ranks 2 and 2, respectively). It is of considerably greater interest to boys than to girls (rank 3 as compared with rank 10).³ *Health* at rank 2 as a problem in 1935 is no longer seen as such (rank 12), though interest in it is still relatively high, especially among boys. *Study habits* are somewhat more of a problem now than in 1935, especially for boys, but the topic ranked quite low in interest value in both periods. *Moral qualities* and *philosophy of life* come next in position both as problems and as topics of interest and are of similar relative magnitude in both periods.

Of intermediate concern as problems are the following: Both sexes see *mental health* as somewhat more of a problem to them now than they did in 1935, and for girls it now appears in the top five ranks. It has, likewise, moved up significantly ($p = .01$) in rank of interest to girls, though it remains low on the boys' list. *Home and family relations* is likewise ranked higher as a problem by girls than by boys now ($p = .01$). Relative to 1935, boys rank it slightly lower and girls slightly higher as a problem, though in these intermediate positions such shifting is statistically meaningless. Both boys and girls rank this topic higher on the list of interests than they did in 1935, girls very considerably so (rank of 2.5 compared with 9, $p = .01$). *Manners and courtesy*, up somewhat ($p = .01$) as a problem for boys as compared with the earlier period, is significantly less a problem for girls now (rank 11 compared with rank 6, $p = .01$). As a topic of interest it is down sharply for both sexes. *Attractiveness*, of intermediate value both as a problem and as an interest for boys in both periods, ranks high in both interest value and as a problem for girls in both periods. Likewise, *getting along with others* was and is of considerable interest to girls and has risen (rank 4 compared to rank 9, $p = .01$) as a topic of interest to boys. It is of intermediate significance as a problem. *Civic affairs* was and is of about median significance as a problem to both boys and girls and is of even less relative importance as a topic of interest.

Of least concern as problems now are the following: *Recreation*, of intermediate significance as a problem to both boys and girls in 1935, is negligible now. It was and is number 1 in interest value to boys but has dropped from second to eighth place for girls ($p = .01$). *Health*, of high significance both as a problem and as a topic of interest to boys and girls in 1935, is now negligible as a problem to either sex. It is now only of intermediate interest to girls, though it remains high on the boys' list. *Safety* is of little interest and even less a problem to both boys and girls now than in 1935 ($p = .01$). *Love and marriage*, ranked low as a problem by boys in both periods, has risen to an intermediate rank as a problem for girls. Both sexes now give it an intermediate interest rating, some several ranks higher

³ In 1935 Symonds attributed the high ranks accorded money to the current economic stress. Studies of allowances show that modern youth "never had it so good" from an economic point of view, but inflation and rising standards of living expectancy keep the 'teen age group keenly aware of the medium of exchange.

CHILD DEVELOPMENT

than 22 years ago. The change in both problem and interest values of this topic is statistically significant, highly so for girls. *Daily schedule* is of little significance then or now, both to boys and girls, as problem or as interest.

This observation concerning the specific rankings accorded the issues should be made. Both psychologically and statistically, the highest and lowest ranks are most differentiated in any ranking or scaling procedure. The method in this study called for the identification of the first three positions, and then positions 13, 14, and 15. The intermediate positions were assigned last and may, particularly in the ranking of interests, represent a state of psychological indifference or lack of discrimination more than a state of intermediate significance. Should this be the case, the significance attributed by Symonds to the "relatively high" rank of philosophy of life (6th in 15 issues) may be modified somewhat. He affirmed (1) that "values and goals are craved" by youth and challenged teacher and counselors by the rank accorded this topic. Symonds attributed the high rank of money as a problem to the depression years and observed "it is a pity" that money drops to a much lower rank in interest value. The problem significance of money was not just a function of depression years, it is now clear.

The lack of concern with love and marriage, identified by Symonds as "sex adjustments," puzzled him, and he explained the low ranks by reference to "repression." His own hypothesis concerning social change appears from the present data to be equally plausible. A similar point can be made about mental health as an issue in adolescence. Symonds dismissed the low ranks attributed to this issue in these words: "Mental health likewise is no concern of healthy, growing adolescents. The crest of life is before them. Their failures and thwartings have not yet turned them back upon themselves" (1, p. 517). The changed cultural ethos apparently has made a difference in the significance accorded this issue in the adolescent years.

The shifts noted in the tables and in the brief discussion for the most part confirm what observers of recent social trends have noted. Today, youth marry younger and show an earlier interest in social relations, love, and marriage. Our culture appears to recognize more openly now than two decades ago the sex, love, and marriage problems of young people. Physical health is actually less a problem today, and possibly receives less attention in school and in the popular press, whereas mental health discussions, literature, and posters appear in every newspaper, magazine, and waiting room. An increase in informality and casualness in dress and behavior may reflect itself in the decline in concern with manners.

The student of adolescent behavior will not be surprised at the significance of money as a problem, high interest in recreation, lack of concern with and interest in safety, unconcern over daily schedule and civic affairs, considerable concern over study habits as a problem but lack of interest in the topic, nor will he be surprised by the greater interest of girls than boys

DALE B. HARRIS

in attractiveness, love and marriage, mental health, and philosophy and beliefs; of boys than girls in money, health, and recreation.

If adults wish to "view with alarm," they may attend to the adolescent's relative unconcern with safety and hazard, set over against the 'teen-age driving problem, and the young person's continuing unconcern with civic affairs, set over against the continued increased emphasis on "modern problems," and citizenship in the secondary school's curriculum theory and effort.

REFERENCES

1. SYMONDS, P. M. Life interests and problems of adolescents. *Sch. Rev.*, 1936, 44, 506-518.
2. SYMONDS, P. M. Sex differences in the life problems and interests of adolescents. *Sch. & Soc.*, 1936, 43, 751-752.

RESEARCH ON RUNNING TIME AND PHYSICAL WORK OF CHILDREN UNDER VARIOUS REINFORCEMENT CONDITIONS^{1,2}

C. G. SCREVEN*

University of Wisconsin—Milwaukee

The investigations to be reported here are a part of a research program on motivation in children being developed by the writer at the University of Wisconsin-Milwaukee. In these particular studies, an effort has been made to examine the simple straight runway as an experimental framework for studying amount of reinforcement (incentive) variables in young children age 4 to 7 years.

Many experimental studies of performance variables have indicated that such factors as amount of reinforcement, drive, or arousal conditions have a great deal to do with performance but little to do with learning per se (1, 2, 7, 8). The factors which do relate to such performance characteristics often are tagged as "performance variables," sometimes as "motivational" variables. These may include the various operations defining drive, emotion, mobilization systems, as well as the amount of reinforcement (Hull's Incentive Motivation, or K). A common method of studying such performance variables is to use the practice levels of well learned responses. One such measure, successfully used in animal studies, is the time taken to

* Department of Psychology, University of Wisconsin—Milwaukee, Milwaukee 11.

¹ This report has been modified from a paper presented on the symposium "Recent Developments in Experimental Methods with Children" at the meetings of the American Psychological Association, Washington, D.C., 1958.

² This research was supported in part by grants from the Graduate Research Committee of the University of Wisconsin. The author wishes to thank Dr. Stanley S. Stahl, Director of the Campus Elementary School of UW-M, and its teachers for their very helpful cooperation. Also, special thanks are due Miss Jeanne Cejka and Mrs. Karen Johnson for their help in running some of the children and to John Porter and Theodore Theander who were most helpful in the design and construction of the apparatus.

CHILD DEVELOPMENT

traverse a straight runway. It is a relatively simple measure. All that is required is that the animal gets from one end to the other. Some kind of learning takes place during the early trials, but the animals soon reach stable speeds which depend upon incentive and drive variables.

About two years ago, pilot studies of kindergarten children's performance in a 14 ft. runway for such incentives as candy, pennies, trinkets, or marble tokens indicated that the runway method might be useful with children. Ss were either given reinforcement on all trials ("high incentive") or on none of the trials ("low incentive"). (Ss ran against an experimentally induced work load to be discussed later.) The children readily adapted to the procedure with a minimum of instruction and, after a few trials, with a minimum of experimenter interaction. Most of the children, when reinforced, developed stable speeds (trial to trial and day to day) and mean speeds were higher under reinforcement as compared to nonreinforcement.

With this encouragement, a more elaborate runway was constructed. Improvements have been made since this time, and, unless necessary, only the latest form of the apparatus will be discussed here.

In addition to the usual measures of reaction time (to starting signal) and running time, it was decided to provide for the control of the amount of physical resistance (or work load) encountered by *S* in traversing the runway. This would allow not only for the examination of physical work as an independent variable but also would provide a gross measure of the amount of work or "energy" (ft. lb. / sec.) which the child expended in traversing the runway.

The reasons for including work may need some additional comment. While it generally is assumed that more "effort" will be made for greater incentives, the experimental facts concerning the relation of physical effort to amount of reinforcement leaves much to be desired. Years ago, Troland (6) talked of motivational (incentive) factors in terms of the physical resistance that would be overcome. This has led to such experimental procedures as the obstruction box approach to the comparison of incentives, where "obstruction" has meant electric shock (not work expenditure). Most of the work and fatigue studies have involved the ergograph procedure, adequately summarized elsewhere (7). These studies have provided much evidence concerning fatigue and work, including the fact that incentive conditions such as knowledge of results, etc., improve performance. But these data do not constitute an examination of the incentive-work relationship implied by many statements about effort and motivation. Perhaps, as Troland suggested, an incentive value can usefully be defined as the amount of physical resistance overcome to achieve it. But, outside of the obstruction box, it is not always clear what we might mean, operationally, by "overcome." In the pioneer study by Fletcher (3), chimpanzees were required to pull a visible banana through an 18 ft. distance against physical resistance (weights). It was found that the larger the reward the more weight was pulled.

C. G. SCREVEN

In the studies reported here, the runway was adapted to investigate further the Fletcher type of problem with children along with the more traditional measures of running speed and reaction time. To accomplish this, a pulley and weight system was constructed. A side view of the main parts of the runway system is shown in Figure 1. The child, in traversing the runway, wears a simple harness attached via a pulley and drum system to a box containing the desired weight. The weights are pulled vertically as the child moves down the runway. Physical work load may be computed in terms of the weight pulled \times distance; the rate of work is equal to the weight \times distance per unit time (sec.).

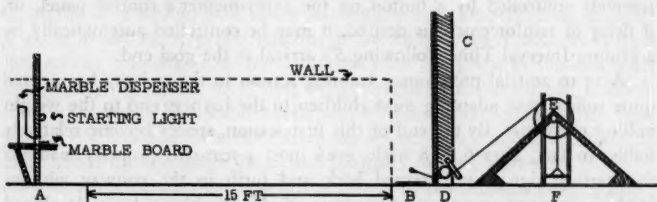


FIGURE 1—Side view of runway showing pulley and weight arrangements.

The clocked length of the runway is exactly 15 ft. At the goal end (A) is a wooden panel containing stimulus lights, a 15 unit marble board and marble payoff compartment on a small ledge, and an automatic marble dispenser. The runway is enclosed by the wall of the room on one side and 5 ft. high panels on the other side. At the starting end is a starting plate (B) and another panel (C) separating *S* from the weight control system and *E*'s control board at the rear. All necessary recording and controlling of stimulus and reinforcing events can be accomplished out of the sight of *S* at this control board. Weights must be changed manually, but this can be done simply and without *S*'s knowledge.

A rope attaches to *S*'s harness and passes through the rear panel (C) and the pulley (D) and onto a large drum (E) as shown in Figure 1. A box (F) containing the weights is attached by a rope to the smaller axis of the drum in such a way that it winds up as the larger drum (E) is turned by *S*'s movement down the runway. The diameters of the drum and axis are such that 5 lbs. of weight in the weight box will provide 1 lb. of work load to *S* (plus minor friction effects).

Three Model S-1 Standard Electric Chronoscopes are used for the timing operations. One timer records reaction (starting) time; the second timer records the time taken to traverse the 15 ft. distance of the runway from the moment *S* leaves the starting plate; the third timer records the time elapsing since the completion of the previous run. This last clock is used to time the interval between successive trials. All timing operations

CHILD DEVELOPMENT

are automatic except for the initial signal which starts the reaction time clock and the signal light to which *S* is to respond. The system is so wired that the running time clock begins only after *S* begins to move and it stops when he reaches a point 15 ft. from the starting end. The latter is accomplished when a clip attached to the rope reaches the lever (*G*) (Figure 1) which opens a holding circuit controlling the clock. To make sure that *S* moves far enough to operate this lever before he stops, the goal panel is placed so that he must stretch slightly to reach the marble (or candy) and marble board. Also, the 15-watt light on the goal panel which serves as the signal to start remains on for the running period, and *S* is carefully instructed not to stop until this light goes off. Marble reinforcement may be remotely controlled by a button on the experimenter's control panel, or, if delay of reinforcement is desired, it may be controlled automatically by a Hunter Interval Timer following *S*'s arrival at the goal end.

A 15 to 20 trial preliminary training session in the runway has proved quite sufficient in adapting most children to the harness and to the weight pulling procedure. By the end of this first session, speeds become relatively stable. In fact, after 6 to 8 trials, even most 4-year-olds properly wait for the starting signal and proceed back and forth in the runway without further conversation or attention from *E*. No problems have developed with the harness and weight pulling arrangement and children 4 years old and older quickly take to the procedure as routine without resistance. Little seems to be gained in procedural smoothness or speed stability by extending preliminary training beyond one session.

A few details of our usual preliminary procedure may be of interest. Immediately upon entering the runway for the first time, the harness is fitted and fastened while *S* is told to watch the light (starting signal) at the other end of the runway (the word "runway," of course, is never used). Nothing is said concerning the specific function of the harness and the weights are not mentioned. No preliminary verbal instructions are given. Each step by step event is allowed to occur in sequence, the proper action then being described. Thus, when the starting light goes on for the first time, *S* is told that this means he can go down to the other end. *E* walks with him on this first usually hesitant trip. (Of course, no references to speed, walking, or running are made.) A 5 lb. work load is used. The type of reinforcement will of course depend upon the experiment, but in our work so far marble rewards have been found quite satisfactory.³ When *S*

³ Marbles rather than candy were used for several reasons. In the first place, some parents object to candy and the accumulated amount of candy with as many as 28 reinforced trials per day could lead to problems. (This may be minimized by using sugared breakfast cereals.) Also, it has been our experience that more than a few children do not want candy or will not eat it on the spot. Finally, the earlier pilot studies involving candy vs. marble reinforcement did not indicate that candy was more effective than the marble tokens. In fact, the effectiveness of marble tokens in relation to other incentives probably is greatly influenced by the details of the marble collecting arrangement itself (instructions, frequency of reinforcement, previous experiences of nonmarble reinforcement, etc.).

receives his first marble, he is shown how to place it on the marble board. He then is told that he can get more marbles in the same way and is taken back to the starting plate. Preliminary sessions involved in the studies reported here have involved a choice of one cat's-eye marble as a take-home prize as the marble board is completed, although a take-home prize probably is unnecessary on this session. The take-home prize is offered when *S* returns to the starting plate following the first trial. (On regular experimental sessions, it is offered at the beginning of the session.) Instructions are repeated as each event occurs on the second trial. On the third trial no marble reinforcement is given. As *S* waits for the marble to come out, *E* explains that sometimes nothing comes out and in this case he should touch the table and go back to the starting plate. By the fifth trial, *E* leaves *S* alone in the runway and remains at the control panel without further interaction. A total of four nonreinforced trials are given in random order (beginning with third trial) and 15 reinforced trials which are sufficient for *S* to complete his marble board before returning to school. Upon completing the board, *S* is released from his harness, his marbles are counted, he is given his take-home prize and returned to his classroom.

We may now turn to some of the results of our runway investigations.

One study, using 18 first grade children as *Ss*, compared reaction time and running scores under relatively high and relatively low amounts of reinforcement per trial. High reinforcement consisted of receiving one marble on each trial; low reinforcement consisted of not receiving any marbles on any trial. *Ss* were assigned at random to the two conditions. The experimental conditions were repeated for each group for a total of three sessions in order to observe day to day stability and possible cumulative effects. *Ss* were given 14 trials on the first session and 28 trials on the second and third sessions. Each *S* was offered his choice from a group of cat's-eye marbles if he filled his marble board. Trials were spaced 30 sec. apart and all *Ss* pulled against a 5 lb. work load.

The main results are shown in Figure 2. Scores are *speed* scores plotted in terms of the reciprocals of runway time for each trial of the three sessions. The mean speed for the higher incentive (reinforced) group increased significantly over the speed of the lower incentive (nonreinforced) group during each of the three sessions. A group \times trials analysis of the final 10 trials of each session indicated that these differences (between groups *MS*) were statistically significant in each case beyond the .001 level ($F = 45.9, 91.4, \text{ and } 73.7$, respectively; $df = 1 \text{ and } 120$.)

As in the pilot studies, the reinforced *Ss* individually exhibited remarkably little variation in speed (frequently no more than $\frac{1}{2}$ sec. for the 15 ft. distance over all 70 trials of the three sessions). Nor was variability greatly increased by nonreinforcement. This is in contrast to the reported effects of failure and nonreward on the variability of other response measures such as speed of crank turning (5). Seven *Ss* previously had participated as kindergarten children in the first runway study over one year earlier, and

CHILD DEVELOPMENT

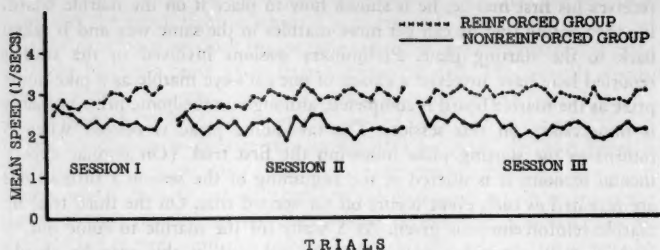


FIGURE 2—The effect of marble reinforcement and nonreinforcement on speed of running.

when a comparison was made of their speeds, it was found that all seven Ss maintained the same ranks over the one year period. The same held true for the variability characteristics of speed. This is a very small sample for generalization but it seems a noteworthy result, and the study of subject (personality) variables which may relate to performance characteristics in the runway might be interesting.

An analysis of the reaction (starting) time scores for the three sessions indicated that, with successive sessions, reaction time was the same for both groups on the first session, remained constant for all three sessions for the reinforced group, but decreased for the nonreinforced group (total mean drop = 0.3 sec.). The reliability of this slight but steady cumulative depressing effect of nonreinforcement on mean reaction time is supported by a group \times sessions trend analysis (4) which yielded an interaction MS significant between the .025 and .05 level ($F = 4.22$; $df = 2$ and 24).

The first study in which physical work was varied involved 17 kindergarten children, all experiencing relatively high reinforcement. Physical work, as mentioned earlier, is generally regarded as having an inhibitory effect on performance. But greater work is done for more incentive, as shown in the Fletcher study, in the ergograph studies (7) and others. One question is what happens if incentive level remains the same while work load or required effort increases. This first study concerned this question and set out to determine how running speed would change with increasing work, incentive conditions constant.

Ss were run for 10 trials a day for seven days, half under increasing work load and the other half under a constant 1 lb. work load. Both groups ran under 1 lb. for the first two days. Control Ss continued at this 1 lb. load throughout the seven day period. Experimental Ss were increased to 2½ lbs. on the 3rd and 4th days, 3¾ lbs. for the 5th and 6th days, and 5 lbs. on the 7th day (the maximum which the first work control apparatus could safely take). All Ss received candy, small toys, or marble reinforcement on eight out of 10 trials per day (80 per cent reinforcement) with

two nonreinforced trials at random points. Types of candy were varied from trial to trial and day to day. Ss received two marbles each day which were collected on a 14 unit marble board. Thus, all seven days of the experiment were required to complete it. When filled, Ss could turn in the marbles and choose from a large variety of take-home prizes on exhibit in Christmas fashion on a table near the runway. These elaborate reinforcement and prize arrangements were used in the hope of maintaining high incentive level throughout the seven day period. However, the marbles-taken alone probably would have been quite adequate for this purpose.

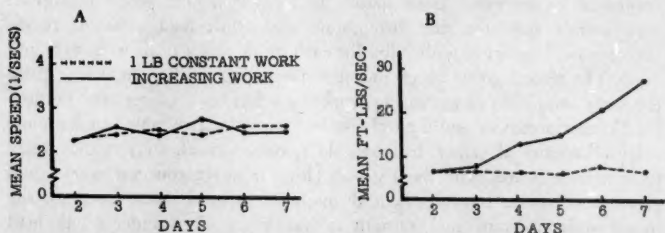


FIGURE 3—The effect of increasing work load on running speed and rate of work output under high reinforcement.

The results of this study are shown in Figure 3A in terms of speed scores. No reaction time data were taken in this experiment. As may easily be seen, running speed did not change with increased work load but remained almost exactly the same as the speed of the control Ss. This was accomplished, theoretically at least, at the expense of doing more physical "work." For example, the rate at which work was done by both groups on days 2 through 7 is shown in Figure 3B where time scores have been converted into ft.-lb./sec.

Several interpretations present themselves. The children may have learned from previous experience that effort against resistance is more often reinforced than a lessening of effort. Also, increased work could, at lower levels, increase drive strength enough to cancel inhibitory effects. In terms of neurological mechanisms, it may be that incentives in some way cancel fatigue effects (for example, by canceling the inhibitory sensory feedback from greater work). In this case, fatigue or work would not have any observable effect until it becomes equal or greater than the total incentive stimulation present in the situation. Ergograph studies as well as other investigations have indicated that the inhibitory effects of fatigue probably are largely controlled by central rather than local processes in the nervous system. Hence, there is the possibility that central neural activities associated with incentive or reinforcing stimulation could compete with and inhibit the effects of inhibitory sensory feedback from muscles. This would

CHILD DEVELOPMENT

suggest not only that the amount of work that will produce a decrement in output depends upon incentive conditions but also that the effects of incentive stimulation on performance will, at least in part, depend upon the amount of work. That is, lower work load will not distinguish performance under different incentive levels as clearly as will higher work load.

Our second work experiment concerned this latter possibility. The experiment, using 24 first grade children, compared the effects of increasing work loads under both high and low incentive conditions. After preliminary training of the type already described, Ss were assigned at random to three groups of eight Ss each. All data were obtained during a single session of 30 successive trials spaced 30 sec. apart. One group was given zero marble reinforcement throughout, and work loads were increased successively, beginning with 1 lb., for each block of six trials to 2, 5, 8, and 10 lb. The second group (high incentive increasing work group) was given the same conditions of increasing work loads but with 50 per cent random marble reinforcement and their choice from an attractive table of take-home prizes (varieties of candy, balloons, dolls, cars, figures, etc.) if they filled their marble board. The third group (low incentive constant work) was given zero marble reinforcement throughout, as in the first case, but the initial work load remained constant so that they worked under a 1 lb. load throughout the session. The nonreinforcement groups were offered a cat's-eye marble if they filled their marble board, but they were not shown the table of prizes used for the reinforced group.

Figure 4 shows the mean running times per trial for each of the three groups. The low incentive increasing work group (LIW) sharply increases

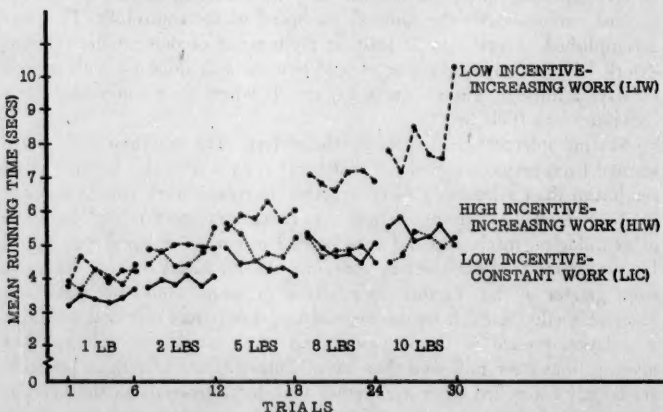


FIGURE 4—The effect of increasing work load on running time under high and low incentive conditions.

its running time (i.e., slows down) as work increases; the other two groups gradually increase a small amount in time scores. No important differences in performance developed between any of the groups before the 5 lb. level. Also, note that no differences developed between the HIW group and the LIC Ss who experienced a constant 1 lb. work load throughout the session. A groups \times work levels trend analysis (4) of running times covering the five weight loads for the three groups yielded a groups \times work levels interaction MS significant beyond the .001 level ($F = 8.08$; $df = 8$ and 80), thus indicating the dependence of work level effects on incentive level and vice versa. In other words, running time does not distinguish between different incentive levels at low work loads but only at higher work loads, apparently the higher the better. High incentive Ss did decline somewhat in their speed as work load increased, but the same decline also occurred for LIC Ss.

According to the results obtained for the 1 and 2 lb. condition, it appears that, had we used lighter weights or ran children in a normal runway without weights, we would not have obtained a difference between the reinforcing conditions used. This is not to say that different incentive conditions cannot be distinguished using the runway in the normal manner, although we have not tried it with children. Certainly it has worked with rats. But, on the basis of these studies, it would appear that it would require greater extremes of reinforcement to do it. What the use of work loads appears to do, if these results are to be taken seriously, is to increase the sensitivity of runway performance as a measure of incentive differences.

In conclusion, the lack of differences at lower work levels and the continued resistance of the reinforced group to increasing work loads is consistent with the possibility, described earlier, that incentive stimulation may compete with and effectively inhibit work or fatigue effects, perhaps until the latter are equal to the total value of incentive stimulation. If so, an operational definition of incentive values in children might be based (after Troland) upon the amount of work load required to produce a decrement in speed. But, however appropriate or inappropriate such speculations, these results do seem to indicate that the effort variable may have to be given more consideration in the study of performance than in the recent past.

REFERENCES

1. BUGELSKI, B. R. *The psychology of learning*. New York: Holt, 1956.
2. CRESPI, L. P. Amount of incentive and level of performance. *Psychol. Rev.*, 1944, 51, 341-357.
3. FLETCHER, F. M. Effects of quantitative variation of food incentive on the performance of physical work by chimpanzees. *Comp. Psychol. Monogr.*, 1940, 16, 1-47.

CHILD DEVELOPMENT

4. LINDQUIST, E. F. *Design and analysis of experiments in psychology and education*. New York: Houghton-Mifflin, 1953.
5. SCREVEN, C. G. The effects of nonreward and interference on the variation in the amplitude of an instrumental response. *J. comp. physiol. Psychol.*, 1955, 48, 299-304.
6. TROLAND, L. T. *The fundamentals of human motivation*. New York: Van Nostrand, 1928.
7. WOODWORTH, R. S., & SCHLOSBERG, H. *Experimental psychology*. New York: Holt, 1954.
8. ZEAMAN, D. Response latency as a function of the amount of reinforcement. *J. exp. Psychol.*, 1949, 39, 466-483.

SIMILARITY OF STIMULI AND OF RESPONSES IN THE SUCCESSIVE DISCRIMINATION PROBLEM

CHARLES C. SPIKER* and RUTH B. HOLTON¹

State University of Iowa

In a recent article Spence (2) analyzes the conventional successive discrimination problem as it is used with rats. He suggests that the problem be viewed as one involving four cue-position patterns (e.g., black-left, black-right, white-left, and white-right), two of which are presented on each trial. The *S* is reinforced for approaching one of these patterns on each trial and is nonreinforced for approaching the other. Spence's analysis suggests a similar one for a problem that has been used with human *Ss* (3). In the latter problem *S* is required to associate each of several stimulus lights with a different one of several response buttons. The stimuli are presented singly and *S* chooses among the available response buttons. It is possible, then, to specify light-button patterns that are analogous to the cue-position patterns. When a given stimulus light (S_A) is presented, there are as many light-button patterns from which to choose as there are buttons; that is, S_AS_{B1} , S_AS_{B2} . . . S_AS_{Bk} . The *S* is reinforced for selecting one of these patterns on each trial and is nonreinforced if any of the others are selected. Viewed in this way, it is apparent that the similarities among these patterns can be manipulated by varying either the similarity of the lights or that of the buttons.

In a previous study by the first author (3) preschool children performed on a light-button task in which there were more response buttons than stimulus lights. The response buttons which were correct for the stimuli may be called the intraset buttons; those which were not correct for any stimulus may be called the extraset buttons. Two types of error measures were used. An *intralist error* resulted from the subject's pushing of an incorrect intraset button; an *extralist error* was defined as the pushing of any extraset button. It was found that an increase in the similarity among the

* Iowa Child Welfare Research Station, State University of Iowa, Iowa City.

¹ The authors collaborated in the design and conduct of a series of experiments in motor learning. The second author's death occurred prior to the preparation of this manuscript. The first author is responsible for the report in its present form.

CHILD DEVELOPMENT

stimulus lights resulted in a relatively greater increase in intralist than in extralist errors.

The present paper reports two experiments designed to extend knowledge of the manner in which similarity among the light-button patterns affects performance on such a task. In the first experiment similarities among the light-button patterns were manipulated by varying the similarity of the stimulus lights. The effect of this variable on intralist and extralist errors was observed. In the second experiment similarities among the patterns were manipulated by making the extraset buttons visually distinct from the intraset buttons. The effect of this variation on the two types of errors was also noted. It was expected that decrease in the similarity among the stimulus lights would result in a decrease in the intralist errors, while decrease in the similarity between the extraset buttons and the intraset buttons would result in a decrease in extralist errors.

EXPERIMENT I

Method

Apparatus. The apparatus was designed to present, one at a time, four stimulus lights differing in brightness, to each of which *S* responded by pushing one of seven buttons arranged in a semicircle on the response panel. Each of four buttons was correct for a different light, and three were incorrect for all lights. Pushing any button during the presentation of a given stimulus resulted in the immediate onset of a small informational lamp beside the button which was correct for that stimulus. Thus, whether *S* pushed the correct or an incorrect button on each stimulus presentation, he was immediately informed of the button that was correct for that stimulus.

The apparatus consisted of three parts: a response, a stimulus, and a control unit. The *response unit* was a black panel 12 in. \times 20 in. mounted on supports which caused it to slope downward toward *S*. Arranged in a semicircle were seven push-button switches, approximately 3 in. apart. Starting at the left, the switches were numbered from one to seven with small white numerals painted on the panel. Immediately beside each switch was a 6-v. pilot (informational) lamp mounted with the glazed tip projecting through the top surface of the panel.

The *stimulus unit* was a black box 6 in. wide, 5 in. high, and 8 in. deep. Centered on the front face of the box was a 2½ in. flashed opal glass aperture. Inside the box, approximately 6 in. behind the aperture, a 120-v., 75-w. projection bulb was mounted so that the beam was directed on the aperture. During the experiment, the stimulus unit was centered immediately above the back edge of the response unit in plain view of *S*.

The *control unit* contained four variable transformers (Powerstat, Type 10), which were used to vary the voltage on the stimulus light circuit, and a jack and plug system which permitted *E* to connect any given button

(and its corresponding informational lamp) with a given stimulus. A selector switch permitted *E* on each presentation to turn on the appropriate stimulus and simultaneously to select the appropriate button and informational light.

Subjects and design. The *Ss* were fourth-grade children obtained from the Jefferson Elementary School of the Ottumwa Public School System.² Three groups of 17 *Ss* each were used. The conditions for the three groups differed only in the similarity of the stimulus lights which were presented. Group D received stimuli that were obtained by setting the four variable transformers to impress 120, 90, 65, and 55 volts on the circuit to the stimulus lamp. Group S received stimuli obtained with 110, 87, 72, and 60 volts. Group HS received stimuli obtained with 105, 95, 85, and 75 volts. This procedure resulted in stimuli varying in brightness and, to some degree, in hue, since the dimmer stimuli also appeared yellow and yellow-orange.

In each group the brightest stimulus (B) was always associated with the first button, the next bright (NB) with the fifth button, the next dim (ND) with the seventh button, and the dimmest (D) with the third button. Thus, the second, fourth, and sixth buttons were never correct for any stimulus.

Procedure. Upon entering the experimental room, *S* was informed that the experiment was designed to determine whether or not children could distinguish among lights better than adults could. He was told that each time a light was presented he was to push a button to attempt to turn it out. He was informed that following each response, whether or not it was correct, the informational light would tell him which button was the correct one. Although all *Ss* were run in a semidarkened room with blinds over the windows and without artificial illumination, the numbers beside the buttons were clearly visible to *E* and presumably to *S*.

A trial was defined as the successive presentation of all four stimulus lights in some order. Each *S* was given a total of 24 such trials. Six different orders of presenting the stimuli were used within each successive block of six trials. The *S* was given as much time to respond as he chose. The interval between presentations varied, since it depended on *E*'s rate of setting the selector switch and on *S*'s response time, but it averaged about 3 sec.

Two error measures were obtained for each *S*. Selection of an incorrect intraset button was counted as an intralist error. Selection of any extraset button was scored as an extralist error.

Results

The mean number of intralist and extralist errors per trial are presented graphically in Figure 1. It will be noted that the number of intralist errors

² The investigators are indebted to Mr. R. O. Wright, Director of Curriculum, Ottumwa Public School System, Ottumwa, Iowa, for his generous permission to use the *Ss* of these experiments. Thanks are also due to Mr. Cecil J. Stevens, and his teaching staff, for their general assistance and for their fine cooperation in providing facilities.

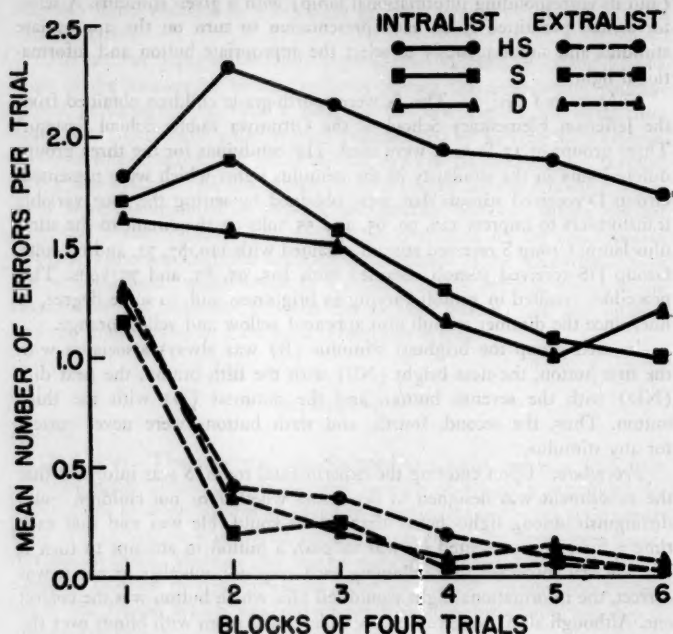


FIGURE 1—Mean number of intralist and extralist errors per trial in blocks of four trials (Experiment I).

was greater than the number of extralist errors for each group, the number of intralist errors was greater the more similar the stimuli, and the frequency of extralist errors did not appear to be much affected by stimulus similarity.

Table 1 presents the means and standard deviations of the total intralist and extralist errors for the entire 24 trials. A summary of the analysis of variance of these data is presented in Table 2. The main effects of stimulus similarity and type of error were found to be significant at a high level of confidence. The interaction between type of error and similarity of stimuli was significant at the .025 level, indicating that with increasing similarity there is a greater difference between the intralist and extralist error means. A simple analysis of variance was conducted separately for the two types of error. Analysis of the extralist errors revealed no significant differences among means of the three groups ($F < 1.00$). Analysis of the intralist errors yielded a significant F ($p < .005$), reflecting the increase in intralist errors with increase in stimulus similarity.

CHARLES C. SPIKER and RUTH B. HOLTON

TABLE 1

MEANS AND STANDARD DEVIATIONS OF TOTAL INTRALIST AND
EXTRALIST ERRORS IN 24 TRIALS
(EXPERIMENT I)

Groups	INTRALIST ERRORS		EXTRALIST ERRORS	
	Mean	SD	Mean	SD
Group D	32.29	13.45	8.47	6.96
Group S	34.24	10.20	6.94	3.84
Group HS	46.53	10.80	9.47	6.65

TABLE 2

ANALYSIS OF VARIANCE OF TOTAL INTRALIST AND EXTRALIST ERRORS
(EXPERIMENT I)

Source	df	Mean Square	F	p
Between Subjects	50			
Stimulus Similarity (S)	2	641.37	8.56	.001
Error (b)	48	74.95		
Within Subjects	51			
Error Type (E)	1	22031.43	245.91	.001
E × S	2	399.30	4.46	.025
Error (w)	48	89.59		

Figure 1 suggests that the magnitude of increase in intralist errors from the first to the second block of trials is an increasing function of stimulus similarity. A sensitive test of this hypothesis can be obtained by using the intralist errors on blocks one and two for Groups D and HS only. An analysis of variance, similar to the one summarized in Table 2, resulted in a significant interaction between trial blocks and stimulus similarity ($p < .05$), indicating that the curves for the two groups over the first two blocks of trials are not parallel. Moreover, for Group HS only, the related t test of the difference in mean intralist errors for these two blocks of trials showed the increase to be reliable ($p < .025$).

EXPERIMENT II

Experiment II was designed to study the effects of button similarity on the intralist and extralist errors. Specifically, a white square was placed around each intraset button or around each extraset button, thereby increasing the distinctiveness of the two sets of buttons. It was expected that this

CHILD DEVELOPMENT

procedure would increase the difference between the mean frequencies for the two types of errors.

Method

Apparatus. The apparatus was like that of Experiment I except that eight response buttons were used instead of seven. The use of eight buttons equated the number of intralist and extralist buttons and avoided the possibility of inadvertently giving *S* advance information as to which were the intralist and which were the extralist buttons when the squares were placed on the buttons. Four white plastic 2½ in. squares were used, each with a hole in the center through which the button projected. Thus, even with a square in place, the button it surrounded was readily available to *S*. The four variable transformers were set at 105, 87, 72, and 65 volts, producing stimuli more distinctive than those of Group HS in Experiment I but more similar than those of Group S.

Subjects and design. Forty *Ss* from the fifth grade² were randomly assigned in equal numbers to an experimental (Group E) and a control group (Group C). To counterbalance for possible button preferences in comparisons of intralist and extralist errors, two subgroups were formed in each major group through the random assignment of half the *Ss* in each. For one subgroup the first button was correct for stimulus B, the fourth for NB, the seventh for ND, and the third for D. For the other subgroup the second button was correct for B, the sixth for NB, the eighth for ND, and the fifth for D.

The white squares were placed on buttons for Group E only. One-half the *Ss* in each subgroup began the task with the white squares placed around the four extralist buttons and the other half of each subgroup began the task with the squares placed around the four intraset buttons. The *Ss* in Group C performed on the task without squares.

Since no significant differences in performance appeared either between the subgroups counterbalancing for button preferences or between subgroups counterbalancing for the placement of squares, these factors have been omitted in the presentation of the results.

Procedure. For Group E, the squares were placed on the appropriate buttons before the *Ss* were admitted to the experimental room. The procedure in Experiment II was otherwise identical to that in Experiment I.

Results

The mean numbers of extralist and intralist errors are graphically presented in Figure 2 in blocks of four trials. The large difference in the frequency of intralist and extralist errors will again be noted. While the two groups do not differ very greatly in number of intralist errors, rather marked differences in extralist errors are apparent; that is, there appears to be an interaction between the groups on the one hand and the type of error on the other.

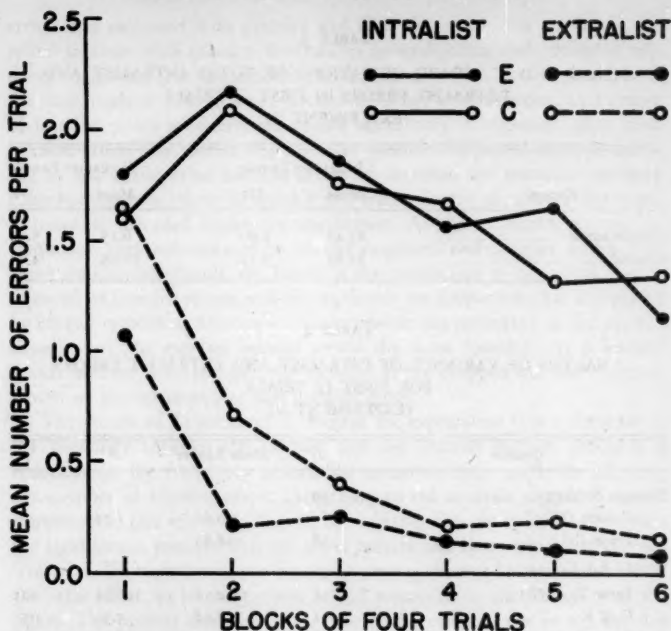


FIGURE 2—Mean number of intralist and extralist errors per trial in blocks of four trials (Experiment II).

The variances for total errors in all 24 trials did not prove to be homogeneous for the different combinations of groups and error types. Therefore, the statistical analyses have been restricted to the first 12 trials, where the assumption of homogeneity of variance was satisfactorily met. This restriction is not considered serious, since changes in the curves are not marked after the first half of the trials, and since the results of analyses of the total data are in close agreement with those of the first half.

Table 3 presents the means and standard deviations of both intralist and extralist errors for each group. Table 4 presents the summary of the analysis of variance of these data. Intralist errors were found to be significantly more frequent than extralist errors, and the interaction of groups and error type was also significant ($p < .05$). A simple analysis of variance of the intralist errors alone showed that they do not differ significantly for the two groups ($F < 1.00$). The same analysis of extralist errors, however, indicates that Group E made significantly fewer errors than did Group C ($p < .005$). Both Groups C and E made significantly more intralist than extralist errors ($p < .001$).

CHILD DEVELOPMENT

TABLE 3

MEANS AND STANDARD DEVIATIONS OF TOTAL INTRALIST AND
EXTRALIST ERRORS IN FIRST 12 TRIALS
(EXPERIMENT II)

<i>Groups</i>	INTRALIST ERRORS		EXTRALIST ERRORS	
	<i>Mean</i>	<i>SD</i>	<i>Mean</i>	<i>SD</i>
Experimental	23.45	7.60	6.15	5.35
Control	21.80	5.33	10.95	4.62

TABLE 4

ANALYSIS OF VARIANCE OF INTRALIST AND EXTRALIST ERRORS
FOR FIRST 12 TRIALS
(EXPERIMENT II)

<i>Source</i>	<i>df</i>	<i>Mean Square</i>	<i>F</i>	<i>p</i>
Between Subjects	39			
Groups (G)	1	49.61	1.72	
Error (b)	38	28.80		
Within Subjects	40			
Error Type (E)	1	3962.11	100.94	.001
G × E	1	208.02	5.30	.05
Error (w)	38	39.25		

DISCUSSION

The results of Experiment I are in agreement with previous findings (3) that an increase in stimulus similarity results in a relatively greater increase in intralist than in extralist errors. In the previous experiment, however, extralist errors also increased with increased stimulus similarity while in Experiment I there was no significant increase. The procedure of the present experiments, by which *S* is automatically corrected after responding, may well account for this discrepancy. The automatic correction procedure speeds up the rate of learning, thereby reducing the number of failures, and produces more homogeneous rates of learning by the *Ss*. These effects combine to yield stable learning curves which reflect more sensitively the experimentally induced variations in the learning process.

Due to homogeneous learning rates, it was possible to show in Experiment I that the increase in intralist errors from the first to the second block of trials is an increasing function of stimulus similarity. In a study of verbal paired-associate learning Gibson (1) reported that the frequency of intralist

errors first increased with practice and then decreased. She attributed the initial increase with practice to stimulus generalization and presented evidence suggesting that the similarity of the stimuli within the list determined the magnitude of this increase. According to her interpretation, an increase of intralist errors with practice should occur only if responses other than intralist errors and correct responses are possible. In paired-associate learning by the anticipation method errors of omission are common on early trials but drop out relatively quickly as *S* learns the list of relevant responses. Extralist or intruded errors are uncommon. As the omissions decrease in frequency, they are replaced by correct responses and intralist errors. The more similar the stimuli, the larger is the proportion of omissions that is replaced by intralist errors, and the smaller is the proportion that is replaced by correct responses. Although omissions were not permitted in the present experiments, the extraset buttons served the same function. As *S* learned which buttons were relevant, the extralist errors dropped out and were replaced by correct responses and intralist errors.

The results of Experiment II confirm the expectation that a decrease in the similarity between the extraset and the intraset buttons produces a reduction in the frequency of extralist errors without markedly affecting the number of intralist errors. According to the analysis suggested in the introduction, this effect is achieved by reducing the generalization between the light-button patterns that are never correct and those which are correct. Theoretically, variation in the similarity among intraset buttons should have the same effect on intralist errors as did variation in the similarity of the lights. The present data, however, provide no evidence for or against this hypothesis.

SUMMARY

Two experiments were conducted with elementary school children using a successive discrimination problem in which there were response choices (push buttons) available that were not correct for any of the stimuli (lights varying in brightness). In the first experiment three groups performed under varying degrees of stimulus similarity. It was found that increased similarity of the lights resulted in an increase in errors involving buttons that were correct for lights other than the one presented (intralist errors), but no significant change was found for errors involving buttons that were never correct (extralist errors). In the second experiment distinctive markings were placed either on the intraset buttons or on the extraset buttons for the experimental group, while the control group performed on the task without the markings. It was found that Group C averaged significantly more extralist errors than did Group E, but there were no differences between the two groups in the number of intralist errors made. The relationship of these findings to current views of stimulus generalization among patterns was discussed.

CHILD DEVELOPMENT

REFERENCES

1. GIBSON, ELEANOR J. Intralist generalization as a factor in verbal learning. *J. exp. Psychol.*, 1942, 30, 185-200.
2. SPENCE, K. W. The nature of the response in discrimination learning. *Psychol. Rev.*, 1952, 59, 89-93.
3. SPIKER, C. C. Effects of stimulus similarity on discrimination learning. *J. exp. Psychol.*, 1956, 51, 393-395.

AN ITEM ANALYSIS OF THE CHILDREN'S MANIFEST ANXIETY SCALE¹

A. JACK HAFNER*² and ARTHUR M. KAPLAN³

Washington University School of Medicine

The children's manifest anxiety scale (CMAS) has been described in detail by Castaneda *et al.* (1). It is an adaptation for fourth, fifth, and sixth grade children of the Taylor scale of manifest anxiety and is comprised of 42 anxiety items and 11 L (lie) scale items. The original normative study (1) showed girls to score higher than boys on the anxiety items. The effects of grade and interaction between sex and grade were found to be nonsignificant. The L scale results showed girls to score higher than boys and sixth graders to score lower than the other two grades. Intercorrelations between the anxiety and L scale items clustered around zero.

The present study involved a general item analysis of the CMAS. In addition, an item analysis was made on the basis of sex and high and low anxiety levels.

PROCEDURE

The CMAS was administered to 122 fifth grade children attending three public schools in the St. Louis area. The group was comprised of 62 boys and 60 girls with a mean age of 134.5 months and an SD of 7.36. The subjects were drawn primarily from middle class families. There were three Negroes in the group of 122 subjects. High (HA) and low (LA) anxiety groups were determined by selecting the upper and lower 20 per cent of the distribution of scores on the CMAS.

* Department of Psychology, University of Minnesota, Minneapolis 14.

¹ The authors wish to express their appreciation to Shirley Millstone for her assistance.

² Now at the University of Minnesota.

³ Now at Eastern Maine Guidance Center, Bangor, Maine.

CHILD DEVELOPMENT

RESULTS

Tables 1 and 2 present the general findings for the anxiety and L scale. The normative findings of Castaneda *et al.* are also included in these tables for comparative purposes. Because of the apparent differences in findings of the two studies, *t* tests were computed for the anxiety and L scale items comparing the results of the two groups of boys as well as the girls. The anxiety scale mean for the boys of the present study was significantly lower and the L scale mean significantly higher than that found by Castaneda *et al.* for their boys. The differences between the girls of the two studies were in the same direction as those found for the boys, but the differences did not quite reach an acceptable level of significance.

TABLE 1
COMPARISON OF ANXIETY SCALE MEANS AND STANDARD DEVIATIONS
FOR FIFTH GRADERS

	PRESENT STUDY			CASTANEDA <i>et al.</i>			<i>t</i>
	<i>N</i>	<i>M</i>	<i>SD</i>	<i>N</i>	<i>M</i>	<i>SD</i>	
Boys	62	13.10	7.29	71	16.24	7.43	2.44*
Girls	60	14.28	8.54	68	17.75	9.23	1.94

* $p < .02$.

TABLE 2
COMPARISON OF L SCALE MEANS AND STANDARD DEVIATIONS
FOR FIFTH GRADERS

	PRESENT STUDY			CASTANEDA <i>et al.</i>			<i>t</i>
	<i>N</i>	<i>M</i>	<i>SD</i>	<i>N</i>	<i>M</i>	<i>SD</i>	
Boys	62	4.13	2.34	71	1.94	1.97	5.82*
Girls	60	3.93	2.21	68	3.07	2.27	1.90

* $p < .001$.

A tallying of the anxiety items for the total group of 122 Ss showed six items to be answered "Yes" by a large majority (60 per cent or more) of the Ss: 6, 7, 29, 33, 44, and 46. The items least frequently (15 per cent or less) answered "Yes" were 1, 8, 13, 23, 26, 37, and 39. The results are presented in Table 3. On the L scale only item 5 was answered in the lie direction by more than 60 per cent of the total group. Item 49 was the only item that was answered in the lie direction by less than 15 per cent of the total group. These results are presented in Table 4.

In a comparison of the Ss' answers by sex a chi square test was computed for each of the CMAS items. The results are presented in Tables 3 and 4. Three of the 42 anxiety items yielded differences significant at the .05 level or less. The items were 14, 48, and 50. In each of these three items more girls than boys gave "Yes" answers. One of the 11 L scale items significantly differentiated between sexes. This was item 30 with more boys than girls answering in the lie direction.

Of the Ss in the LA group none of the anxiety items were answered "Yes" by a majority (50 per cent or more). However, those anxiety items answered "Yes" most frequently (20 per cent or more) were 6, 7, 27, 29, 33, 44, and 46. All of the anxiety items differentiated between the LA and HA groups in the expected direction. A chi square comparison of LA and HA group answers on the L scale items did not yield any significant differences. The results are presented in Tables 3 and 4.

A comparison of the anxiety items was made between the HA group and the remaining 80 per cent of the total group of 122 Ss. Chi square tests resulted in significant differences ($p < .05$) for 32 of the 42 anxiety items. In all 32 cases the HA group had a higher frequency of "Yes" answers. These results are presented in Table 3.

DISCUSSION

In examining the anxiety items answered "Yes" by 60 per cent or more of the total group of Ss, the wording of item 6, "I notice my heart beats fast sometimes," and item 7, "At times I feel like shouting," may have been important in the items being answered the way they were. These two items are the only anxiety items which use the qualifying terms "sometimes" or "at times" which enable a person to answer the items affirmatively without implying severity of a symptom. One would speculate that most of the anxiety items would be answered affirmatively if they were so qualified since the items refer to behavior experienced by most people at one time or another.

With regard to other anxiety items answered most frequently "Yes," the content of item 29, "I worry about doing the right things," and item 44, "I often do things I wish I had never done," seem to be essentially concerned with the same things. It is not surprising, therefore, that both are answered the same way by a majority of the Ss. These items, along with item 33, "I worry about how well I am doing in school," would seem to reflect some typical middle class concerns. Since most of the Ss were from the middle class, it is understandable why the majority might answer these particular items affirmatively. In regard to item 33, it should be noted that the L scale item least frequently answered in the lie direction was item 49, "It is good to get high grades in school." The majority of Ss answering "Yes" to item 46, "I often worry about what could happen to my parents,"

TABLE 3
ANALYSIS OF ANXIETY ITEMS: "YES" ANSWERS

Anxiety Items	Group:	Total (N=122) %	Boys (N=62)		Girls (N=60)		L.A. (N=26) %	H.A. (N=27) %	H.A. vs. Lower 80% Chi Square
			Freq.	Freq.	Freq.	Chi Square			
1. It is hard for me to keep my mind on anything.		15	9	9		.03	0	48	27.43**
2. I get nervous when someone watches me work.		28	16	18		.10	4	63	19.06**
3. I feel I have to be best in everything.		20	12	12		.02	4	33	3.06
4. I blush easily.		21	11	15		.57	4	48	12.91**
6. I notice my heart beats very fast sometimes.		65	42	37		.02	23	78	1.90
7. At times I feel like shouting.		60	36	37		.05	35	74	2.21
8. I wish I could be very far from here.		13	5	11		1.99	0	37	1.48
9. Others seem to do things easier than I can.		47	31	26		.31	4	85	18.67**
11. I am secretly afraid of a lot of things.		26	14	18		.53	8	63	34.68**
12. I feel that others do not like the way I do things.		33	22	18		.20	4	74	36.15**
13. I feel alone even when there are people around me.		12	8	7		.01	0	37	16.85**
14. I have trouble making up my mind.		44	20	34		6.18*	12	82	17.58**
15. I get nervous when things do not go the right way for me.		38	21	25		.49	0	74	17.59**
16. I worry most of the time.		19	16	7		3.07	0	52	21.99**
18. I worry about what my parents will say to me.		40	24	25		.02	12	67	8.77**
19. Often I have trouble getting my breath.		19	10	13		.61	0	33	3.61
20. I get angry easily.		28	19	15		.24	12	37	.92
22. My hands feel sweaty.		36	24	20		1.85	12	52	2.92
23. I have to go to the toilet more than most people.		6	4	3		.002	0	15	3.35
24. Other children are happier than I.		21	17	8		2.90	0	44	10.39**
25. I worry about what other people think about me.		39	27	21		.61	12	78	19.44**
26. I have trouble swallowing.		5	3	3		.14	0	11	1.40

(continued on next page)

TABLE 3 (continued)
ANALYSIS OF ANXIETY ITEMS: "YES" ANSWERS

Anxiety Items	Group:		Boys		Girls		Boys vs. Girls		LA		HA		HA vs. Lower 80% Chi Square
	Total (N=122)	%	(N=62)	Freq.	(N=60)	Freq.	Chi Square	%	(N=26)	%	(N=27)	%	
27. I have worried about things that did not really make any difference later.	57		37	32	32		.27	31		85		10.12**	
28. My feelings get hurt easily.	28		16	18	18		.10	0		67		23.54**	
29. I worry about doing the right things.	66		43	37	37		.41	46		89		7.08**	
31. I worry about what is going to happen.	48		31	27	27		.14	15		85		17.81**	
32. It is hard for me to go to sleep at night.	21		11	14	14		.25	4		41		7.20**	
33. I worry about how well I am doing in school.	75		47	44	44		.01	42		100		10.21**	
35. My feelings get hurt easily when I am scolded.	43		22	30	30		2.07	8		63		7.94**	
37. I often get lonesome when I am with people.	10		5	7	7		.13	0		33		.15	
38. I feel someone will tell me I do things the wrong way.	35		20	23	23		.26	8		82		29.93**	
39. I am afraid of the dark.	12		4	11	11		2.97	0		26		4.46*	
40. It is hard for me to keep my mind on my school work.	21		12	13	13		.01	0		48		14.17**	
42. Often I feel sick in my stomach.	37		18	27	27		.27	8		74		18.60**	
43. I worry when I go to bed at night.	16		10	10	10		.03	0		44		17.36**	
44. I often do things I wish I had never done.	69		41	43	43		.22	31		89		5.42*	
45. I get headaches.	44		24	30	30		1.51	19		78		14.09**	
46. I often worry about what could happen to my parents.	64		43	35	35		1.16	27		82		4.60*	
48. I get tired easily.	28		9	25	25		9.87**	8		74		33.93**	
50. I have bad dreams.	23		9	19	19		4.15*	0		63		28.55**	
51. I am nervous.	16		10	9	9		.01	0		41		14.34**	
53. I often worry about something bad happening to me.	35		19	24	24		.80	12		74		20.77**	

* $p < .05$.** $p < .01$.

TABLE 4
ANALYSIS OF L SCALE ITEMS

Group:	Total (N = 122) %	Boys (N = 62) Freq.		Girls (N = 60) Freq.		Boys vs. Girls Chi Square		LA (N = 26) %		HA (N = 27) %		LA vs. HA Chi Square
5. I like everyone I know.	63	36	41	.98	20	16						.74
10. I would rather win than lose in a game.	26	14	18	.53	5	7						.94
17. I am always kind.	38	27	19	1.36	14	10						.91
21. I always have good manners.	50	33	28		16	12						.94
30. I am always good.	23	21	7	6.65**	9	6						.34
34. I am always nice to everyone.	40	24	25	.02	12	11						.01
36. I tell the truth every single time.	38	23	23	.002	10	10						.03
41. I never get angry.	41	26	24	.001	10	14						.49
47. I never say things I shouldn't.	43	27	25	.007	9	16						.18
49. It is good to get high grades in school.	7	6	2	1.10	2	1						.001
52. I never lie.	43	27	26	.03	10	10						.02

** $p < .01$.

may reflect some dependency feelings associated with this age group of children. These items (29, 33, 44, 46), nevertheless, were answered "Yes" significantly more often by the HA group than by the rest of the Ss.

It is interesting to note that the anxiety items answered "Yes" most frequently by the total group of Ss were the same items most frequently answered "Yes" by the LA group with the exception of item 27.

The anxiety items admitted to least frequently by the Ss seem to represent a rather heterogeneous grouping, but certain explanations are suggested for some of the items. Two of these items are closely related in regard to content: 13, "I feel alone even when there are people around me"; 37, "I often get lonesome when I am with people." These may not have been admitted to since social inclusion and participation are so important at this age. The infrequent affirmative answering of item 8, "I wish I could be very far from here," may have merely indicated a positive test taking attitude of that particular time in regard to the questionnaire. Item 39, "I am afraid of the dark," was another of the questions most infrequently answered "Yes." This is a fear more frequently associated with a younger age group and admitting to this fear would not seem socially acceptable to the fifth grader.

In considering sex differences on the CMAS, only three of the 42 anxiety items significantly differentiated between boys and girls. On the basis of chance alone a little over two significant differences would be expected in making that many comparisons. The most parsimonious way of interpreting these sex differences, therefore, would seem to be on the basis of chance. The same thing may be said for sex differences on the L scale. For the particular sample of this study, it would seem that the CMAS items do not differentiate on the basis of sex.

The comparison of the present findings with the Iowa group's normative data shows the boys of this study to be significantly more defensive as reflected by a higher L scale mean score. This increased defensiveness may in turn have resulted in the significantly lower mean anxiety scale score as the boys may have been less willing to admit to difficulties. The same trend was also suggested for the girls of the present study. An explanation for the conflicting finding of the two studies may lie in the way in which the CMAS was administered. For the Iowa study the CMAS was administered by the classroom teachers of the children while in the present study the investigators themselves administered the CMAS. Since the present investigators had had no previous contact with the subjects and were introduced to them as being from the University, a general set may have been produced whereby the children were trying to make a favorable impression on the visiting "dignitaries." The result, therefore, may have been an increased defensiveness which may have been particularly marked for the boys as both of the investigators were males. The possibility that conditions under which the CMAS is administered may influence the test results is something that needs further investigation.

CHILD DEVELOPMENT

In examining the anxiety items which were the least discriminating between the HA group and lower 80 per cent of the distribution, it is not surprising to find that items 6 and 7 were poor discriminators since these were the two most qualified anxiety questions as pointed out above. Another poor discriminator was item 20, "I get angry easily." The appropriateness of the content of this item inferring anxiety would seem highly questionable.

There was a group of items pertaining essentially to physiological complaints which were poor discriminators in regard to anxiety level. These were items 6, 19, 22, 23, and 26. The only physiological complaint items that significantly discriminated were items 42, 45, and 48. This means that the greater proportion of physiological complaint items did not prove to be discriminating between the group with the highest anxiety level and the rest of the distribution.

The lack of significant differences between the LA and HA groups on the L scale would seem to give further confirmation to the lack of relationship between the anxiety and L scale items.

SUMMARY

The CMAS was administered to 122 fifth grade children from predominantly middle class families. An item analysis was made of the CMAS and comparisons were made on the basis of sex and high and low anxiety levels. In general, the findings showed that the items did not differentiate between sex groups but that a majority of the anxiety items did differentiate between the high anxiety group and the rest of the distribution. The results were discussed on the basis of the content of the individual items.

REFERENCES

1. CASTANEDA, A., McCANDLESS, B. R., & PALERMO, D. S. The children's form of the manifest anxiety scale. *Child Developm.*, 1956, 27, 317-326.

A RESEARCH NOTE ON FATHER-CHILD RELATIONS AND FATHER VIEWED AS A NEGATIVE FIGURE

THOMAS POFFENBERGER*

University of California

In the fall of 1955 a study to determine factors in the formation of attitudes toward mathematics was undertaken at the University of California at Davis. A 140-item questionnaire was administered to all incoming freshmen, consisting of 390 students, 52 per cent males and 48 per cent females. In order to analyze attitude formation, two extreme groups were selected from the sample: those who had a strong liking for mathematics and those who had a strong dislike of the subject.¹ This was done by selecting all those who, on a five-point scale, checked that they liked mathematics "very much" and also responded to another question that they liked mathematics "better" than any other subject. This group consisted of 68 individuals and was designated as the "like" group. The "dislike" group was made up of those who checked either "dislike some" or "dislike very much" on the five-point scale and also checked that they liked mathematics "less" than other school subjects. The "dislike" group consisted of 75 students. The students were also asked to rate their fathers' attitude toward mathematics by using the first mentioned five-point scale.

One of the most significant relationships found in the study was the student's rating of his own attitude toward mathematics and his rating of his father's attitude toward mathematics.² In accordance with identifi-

* Department of Home Economics, University of California, Davis.

¹ An attempt to hold ability constant was made by using scores on the American Council on Education, Psychological Examination. After selection was made, there was no significant difference between the two groups in over-all high school grades or in their general attitude toward school. Only 3 per cent had received less than a B average in high school.

² Attitude of mother and child toward mathematics was not found to be significant largely perhaps because of the small number of mothers who were indicated as liking the subject.

CHILD DEVELOPMENT

cation theory attitudes of "like" or "dislike" were similar for student and father to a degree significant at the .01 level by the chi square test.

The validity of this relationship was uncertain, however, because the investigators were dependent upon the student's evaluation of the father's attitude. The question was whether the students could give a valid rating of their fathers' attitude toward mathematics. A consideration of identification theory suggested the possibility of a difference in rating depending upon the relationship of the student and his father. A single question on the questionnaire was used to evaluate the psychological distance of student-father relationship. The question asked the student to rate his relationship with his father on a five point scale ranging from "very close" to "very distant." Those who rated their relationship with their fathers as "very close" or "close" were considered the "close" group. Those who rated their relationship with their father as "neither" close or distant, "distant," and "very distant" were considered the "distant" group. Taking only the extreme "like" and "dislike" groups, consisting of 130 students from the total sample of 369 students, their rating of their fathers' attitudes toward mathematics was run against their judgment of their relationship with their father. The hypothesis was that for both those with a distant relationship and for those with a close relationship there should be no significant difference from the population from which they were drawn.

TABLE I

DISTRIBUTION OF THOSE WITH A "CLOSE" AND "DISTANT" RELATIONSHIP WITH FATHER AND THEIR RATING OF FATHERS' ATTITUDES TOWARD MATHEMATICS COMPARED WITH THAT OF THE TOTAL POPULATION

<i>Fathers' Attitude toward Mathematics</i>	<i>Total Population</i>	<i>"Close" Relationship</i>	<i>Total Population</i>	<i>"Distant" Relationship</i>
	PER CENT	PER CENT	PER CENT	PER CENT
"Like"	43	50	43	27
"Dislike"	57	50	57	73
Total N	369	74	369	56
	$\chi^2 = 1.28; p < .30$		$\chi^2 = 5.28; p < .05$	

Table I illustrates that the null hypothesis was accepted where there was a close relationship but was rejected at the .05 level of significance by the chi square test where an admitted distant relationship existed. In the total population 43 per cent reported that their father liked mathematics while 57 per cent reported that their father disliked the subject. Where

a close relationship was expressed, there was no significant difference. Fifty per cent reported that their father liked the subject and 50 per cent reported he disliked it. However, only 27 per cent of those who reported a distant relationship reported their fathers liked mathematics while 73 per cent reported he disliked it. Perhaps students who were not close to their fathers did not actually know them well enough to judge their attitudes. However, whether or not the students knew or did not know their fathers' true feelings, there was a significantly greater tendency for those who did not feel a close relationship to assign a negative attitude to the father. It seems unlikely that these fathers actually varied to any great extent from the total population in their attitude toward mathematics. The explanation for the difference more likely lies in these students' perception of their fathers' attitudes toward mathematics as compared with the perception of those students who had a close relationship. It may be that they tended to take the negative feeling they believed their fathers felt for them and consciously or unconsciously assigned this same feeling to other attitudes of the fathers. In effect, "If father does not like me, father does not like anything."³ Further studies would need to be conducted to see if the same results would be obtained in an analysis of attitudes toward other things.

The findings presented in Table 1 suggest that those students who had a distant relationship perceived their fathers as negatively oriented toward mathematics and perhaps negatively oriented toward things in general. If this is so, the statement on the part of the student that the father either likes or dislikes mathematics then becomes a possible indication of identification with father along with the students' statement of psychological distance. We would then hypothesize that those who reported their fathers liked mathematics and had a distant relationship would be more likely to see themselves similar in attitude to their fathers than would those who stated their fathers disliked mathematics and had a distant relationship. The results of a test of this hypothesis are found in Table 2.

Where the students reported that their fathers liked mathematics, there was no significant difference between those who had a close relationship and those who had a distant relationship with their father in terms of similarity of interest in the subject. In both the "close" and the "distant" groups they tended to have the same attitude as their father. Where the students reported that their fathers disliked mathematics, there was a significant difference between the "close" and "distant" group. Students who reported that their fathers disliked mathematics but claimed a close relationship tended to have similar attitudes. However, the combination of a "distant" relationship with father and the projection of an attitude of dislike toward mathematics, whether real or not, resulted in expressed attitudes of the students that bore no significant relationship with the stated attitudes

³ For the purpose of this paper it is assumed that children perceive a close relationship with their parents as having a positive valence and a distant relationship as having a negative valence.

CHILD DEVELOPMENT

TABLE 2

RELATIONSHIP WITH FATHER, STUDENTS' EXPRESSION OF FATHERS' ATTITUDES TOWARD MATHEMATICS, AND STUDENTS' OWN ATTITUDES TOWARD MATHEMATICS

Students' Attitude	FATHERS' ATTITUDE			
	"Like"		"Dislike"	
	RELATIONSHIP WITH FATHER		RELATIONSHIP WITH FATHER	
	"Close"	"Distant"	"Close"	"Distant"
	PER CENT	PER CENT	PER CENT	PER CENT
"Like"	60	73	22	44
"Dislike"	40	27	78	56
Total N	37	15	37	38
	$\chi^2 = .78; p < .50$		$\chi^2 = 5.01; p < .05$	

NOTE.—Three students in the "Distant" group did not respond to the question on their own attitude toward mathematics and are therefore not included in this analysis.

of the fathers. In spite of a feeling of a "distant" relationship with their fathers, the one group tended to identify, to see themselves as similar in attitude, while the other group did not identify and so did not see themselves as similar in attitude.

CONCLUSION

Students who reported a distant relationship with their fathers perceived them as negatively oriented toward mathematics to a significantly greater extent than did the total population from which the sample was drawn.

Students who reported a close relationship with their fathers did not differ significantly from the total population in their rating of their fathers' attitudes toward mathematics.

Since it is unlikely that fathers who had distant relationships with their children differed significantly from fathers who had close relationships in terms of their attitudes toward mathematics, it is suggested that children who see themselves as negatively perceived by a parent may perceive their parents as being negatively oriented to other aspects of life.

BODY CONFIGURATION AND SCHOOL READINESS^{1,2}

MARIA D. SIMON^{*3}

University of Arkansas Medical Center

Efforts to discover the precise relationship between physical and mental status have hitherto met with only limited success. Evidence is contradictory, ranging from studies that show no consistent relationship between physical and mental indices, such as findings on the relation of physical growth to achievement in adolescence (5, pp. 239-254) or on the relation of brain size to mental productivity (16), to those claiming absolute interdependence of all the growth processes of body and mind which ought to result in close statistical correlations provided certain procedures are followed. Variants of the latter view are put forward by Courtis (4), Millard (17), Olson and Hughes (19), but their more extreme claims still await validation.

* Laboratory for Behavioral Research, Department of Psychiatry, Medical Center, University of Arkansas, Little Rock.

¹ This project was supported by U. S. Public Health Grant M-1091.

² Additional data have been deposited as Document number 6033 with the American Documentation Institute, Auxiliary Publications Project, Photoduplication Service, Library of Congress, Washington, D.C. A copy may be secured by citing the Document number and by remitting \$5.00 for photoprints, or \$2.25 for 35 mm. microfilm. Advance payment is required. Make checks or money orders payable to: Chief, Photoduplication Service, Library of Congress.

³ The author gratefully acknowledges the contributions of Roscoe A Dykman and Charles R. Galbrecht who acted as consultants on all phases of the research, and Judson L. Crow, a sophomore medical student, who carried out the statistical computations. Dr. Dykman is director of the Laboratory for Behavioral Research. Thanks are extended to Mr. Virgil T. Blossom, Superintendent of Little Rock Public Schools, and Mr. and Mrs. A. D. Anthony, Principals of Anthony School, for making the subjects available; Mrs. N. W. Quick and Mrs. C. V. Stabler, Hospital Volunteers, for assistance with the collection of data; and Theo C. Panos, R. S. Ely, Rosalind Abernathy, Alice G. Beard, and W. T. Kniker, pediatricians, who gave generously of their time to act as judges.

CHILD DEVELOPMENT

Most well-documented researches, longitudinal as well as status studies, show a low positive correlation between physical and mental variables, usually between $r = 0.10$ and 0.30 . Outstanding examples can be found in the studies on gifted children by Terman (23) and the Harvard Growth Studies (5).

The mentioned investigations have for the most part dealt with units of absolute growth in an ever increasing number of dimensions such as height, weight, ossification, dentition, carpal growth, strength of grip, IQ, and school achievement. Sometimes several indices were combined, but combination of indices usually yielded no better correlations than did correlations of single dimension indices (22).

In spite of such divergent results, the total picture may not be as confusing as it seems at first glance. From the numerous studies and their different approaches certain areas of agreement seem to emerge. Most students of child development would perhaps agree that: (a) Growth in different dimensions is not always parallel; the individual grows in parts rather than as a unit. For example, a child may at one time be well ahead of his peers with regard to dentition and at the same time be retarded in intelligence. The correlation between two indices at any one time may therefore be low or even inverse. (b) Although growth in different dimensions may not be parallel, there seems to be an underlying growth rhythm in the organism as a whole, resulting in both individual and generalized growth patterns.

Courtis (4) has drawn attention to the cyclical nature of growth, and he distinguishes four over-all growth cycles in the life span of the individual: the cycle of fetal growth; the infant cycle; the cycle of middle childhood or preadolescence; and the adolescence cycle, leading to maturity. The life epochs, as defined by the theory of maturational cycles, accord well with the maturational phases named by a Committee of the American Psychiatric Association (20) which distinguishes the following periods: infancy and early childhood, comprising the period from birth through the fifth year; middle childhood through the ninth year; followed by the phases of puberty and adolescence; and finally adulthood. There is much to suggest that these terms describe fairly accurately genuine biological entities, clearly set off from one another during the process of maturation in the human, and that they are not artifacts of nomenclature. We may be reminded of the fact that in all periods of history, and in practically all existing cultures, children between approximately the ages of 5 to 7 are subjected to what formal instruction the particular culture has to offer ("school-entry age"). The period of adolescence usually marks the end of childhood and schooling and leads to a more or less ceremonious admittance into the adult society.

While adolescence is fairly well recognized as a period of life when a characteristic reorganization of the individual, both in its physical and mental aspects, takes place, leading from middle childhood to adulthood, a minimum of consideration has been given to the psychophysical transfor-

mation attending the transition from early childhood to middle childhood. There is not even a term in the English language describing this period of life which in many respects is just as striking as puberty.

The present study is devoted to an exploration of the nature of these changes—in particular the interrelation of physical and mental maturation—during the uncharted age around school entry. In doing so, the writer has hoped to find a useful new approach to the subject by testing hypotheses first formulated more than two decades ago in papers by Zeller (25) and Hetzer (11). These writers have presented evidence to show that school readiness is associated with changes in body configuration. While Zeller's and Hetzer's papers have found wide and uncritical acceptance in German-speaking pediatric and psychological circles, they are apparently unknown in English-speaking countries. To the writer's knowledge conclusions of these papers have not been checked, although Hammond (8), in England, has reported that children from independent schools surpass children from council schools by about two years in a variety of physical measurements.

Zeller employs the concept of *configuration* (Gestalt, in German) to describe physique, a term which encompasses both the specific relationship of different parts of the body to each other as well as the total form. He holds that the child, at some time between the ages of 5 to 7, undergoes a striking reorganization in bodily form, leading from the physique of early childhood (Kleinkind-form) to the physique of the school child (Schulkind-form). These changes are called the "first change in body configuration," while those of adolescence are called the "second change in body configuration." The presumed psychological correlate of the completed "first transformation" is school readiness.

Figure 1 is a schematic presentation of the three principal body configurations which Zeller distinguishes: the figure of early childhood; the intermediate figure; and the figure of middle childhood or the school child figure.

The early childhood figure is characterized by an appearance of top-heaviness with the head and trunk dominating the extremities; a forehead that is relatively large, rounded, and prominent, while the central and lower parts of the face are delicate, soft, and small; a characteristic facial expression which Zeller calls "uncritical"; a short and stocky neck that merges with a sloping shoulder line; a trunk that is relatively large and sack-like in shape with no apparent waist; a protruding abdomen with the lower thoracic aperture forming a wide angle; narrow shoulders with the shoulders and pelvic regions approximately equal in width; and a body surface that is smooth and taut with the lateral body outlines formed by adipose tissue rather than by muscles and joints.

In the figure of middle childhood the size of the head and the trunk has decreased relative to the extremities; the forehead has decreased in size in relation to the central and lower parts of the face; nose, mouth, and chin are firmly moulded; the facial expression is "critical"; neck length

CHILD DEVELOPMENT

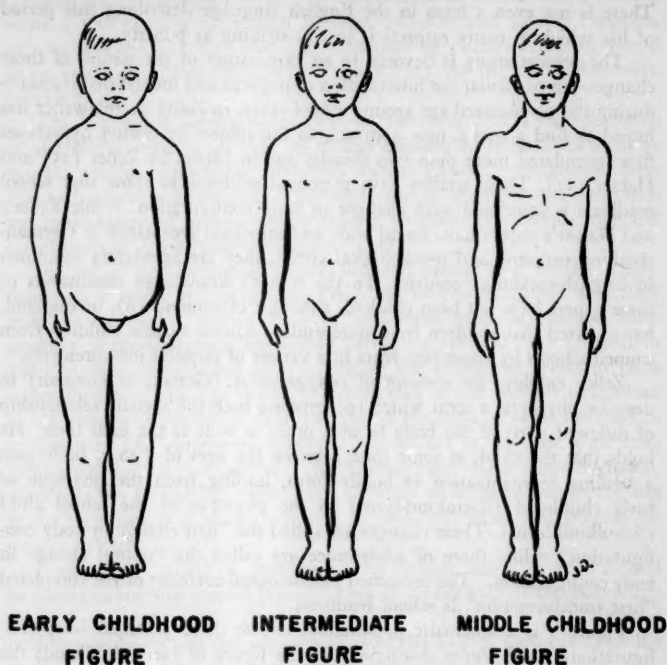


FIGURE 1—The three main body configurations in children of school age.

has increased; the trunk is no longer cylindrical but cone-shaped, broad at the shoulders and narrow at the hips; the shoulder line is square; the infra-sternal angle has narrowed and the abdomen flattened; the waist is clearly indicated; extremities have gained in length and firmness, and the muscles and the joints of arms and legs are clearly visible; the body surface as well as lateral outlines are textured and shaped by increased muscle, while adipose tissue has decreased.

Figures that cannot be grouped easily with either of the foregoing extremes, because they exhibit mixed features, are termed intermediate figures. It is assumed that these children are in a process of transition from the stage of early to that of middle childhood.

Zeller, a physician in the public school health service in Berlin, advised parents of first grade children with the early childhood figure, against

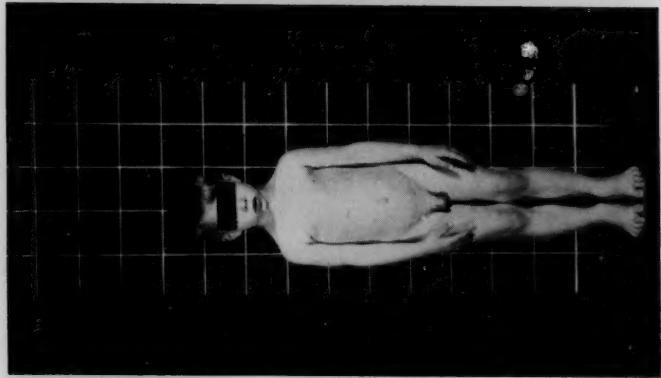


FIGURE 2—M-4, age 4.7. Early childhood figure.

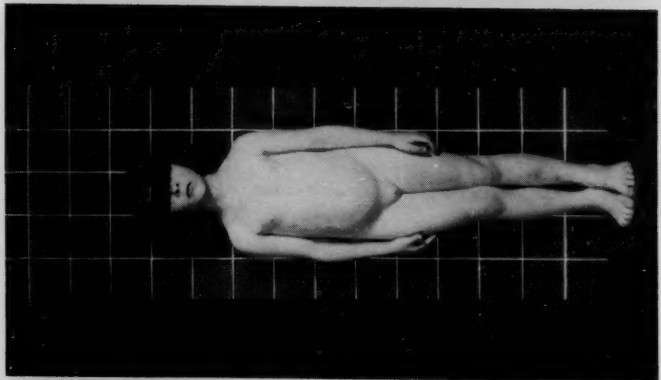


FIGURE 3—F-14, age 5.0. Early childhood figure.

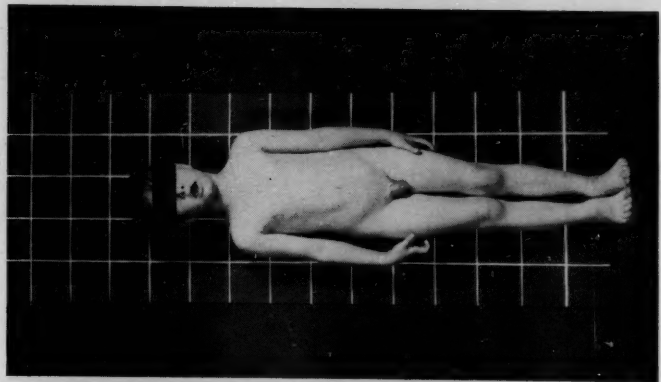


FIGURE 4—M-44, age 5.11. Intermediate figure.

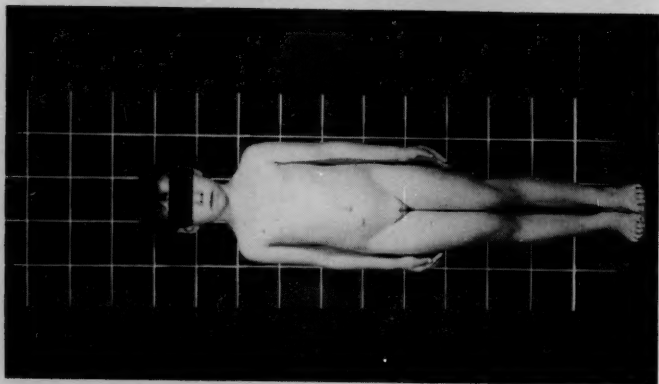


FIGURE 5—F-18, age 5.4. Intermediate figure.



FIGURE 6—M-62, age 6.6. Middle childhood figure.

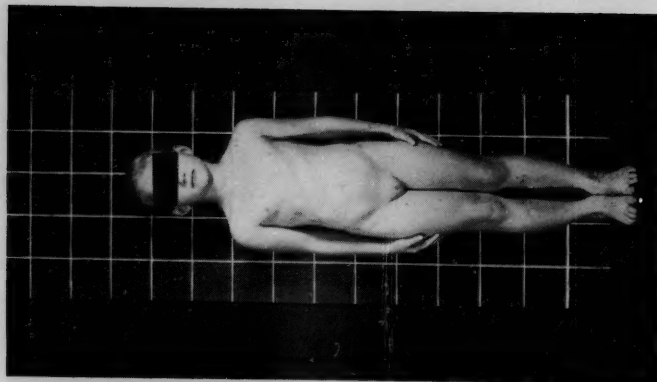


FIGURE 7—F-64, age 6.7. Middle childhood figure.

MARIA D. SIMON

entering these children in school. He justified this on the grounds of general immaturity.

Hetzer tested 40 children between 5 years, 6 months, and 7 years of age, who had been first classified by Zeller into figure types, with the Bühler-Hetzer Developmental Scale. She found that the early childhood figures had the lowest scores and the middle childhood figures the highest, and concluded that children with the early childhood figure, even though of school age, are late developers both physically and mentally.

HYPOTHESES OF THE PRESENT STUDY

Two hypotheses as follows were formulated, the test of the second dependent upon the confirmation of the first.

1. Subjects (Ss) between 5 and 7 years of age undergo a lawful transformation of body configuration identifiable by inspection and measurement.

2. Body configuration is associated with school readiness.

It was necessary to design two separate experiments to test these hypotheses since it was not feasible to gather all the information needed on one population of children.

THE FIRST EXPERIMENT

Procedure

The Ss were 90 Caucasian children, 43 boys and 47 girls, ranging in age from 4-6 to 7-5. There were 30 Ss in each of three age groups, 4-6 to 5-5, 5-6 to 6-5, and 6-6 to 7-5. All Ss in the two younger age groups came from a private school, but only 10 of the Ss in the oldest group were available from this source. The remaining 20 in this group were obtained from the first grade of a public school located in the same neighborhood as the private school. Most of the Ss came from business and professional homes, and a large percentage were children of physicians. These Ss were physically healthy and well nourished.

Body configuration was determined by an elaboration of the methods of Zeller, who employed both the technique of inspection (somatoscopic method) and measurement (somatometric method). In the present study somatoscopy was based on photographs. The 70 Ss drawn from the private school were photographed in the nude, front and side views, against a background grid. The squares of the grid had a side length of 100 mm. One photograph was spoiled, leaving 69 usable pictures, and of these 60 were of Ss 4-5 to 6-5 in age, and only nine were over 6-6.

Five board-certified pediatricians were briefed by the experimenter (E) on Zeller's findings, and were asked to sort (a) all 69 pictures of front views into the three configurations (early childhood, intermediate, middle

CHILD DEVELOPMENT

childhood); (b) 12 pairs of frontal photographs (seven male and five female pairs) of Ss matched for sex and age into a mature and an immature partner; and (c) the 69 pictures of frontal views, with head and genital region concealed, into sex groups. In the second sorting task, the paired pictures were presented twice: once with only the heads exposed and again with only the bodies exposed. Judges were further asked to indicate their reasons for selecting the immature partner.

Somatometric data were obtained for all 90 Ss. They were made on the unclothed body using standard anthropometric procedures (1, 15) and landmarks, unless otherwise noted. Instruments included a regulation anthropometer for heights, spreading calipers for breadths, and a steel tape for girths. The following measurements in mm. were obtained.

1. *Standing height.* From the highest point on the head to the floor (mid-sagittal plane).
2. *Chin height.* From the protuberance of the chin (gnathion) to the floor.
3. *Suprasternal height.* From the middle of the anterior-superior border of the manubrium sterni to the floor.
4. *Leg length.* From the right anterior-superior iliac spine to the floor. Following Zeller, this measure was termed leg length, although on the skeleton, leg length is, as a rule, measured from the trochanter of the femur.
5. *Symphysis height.* From the symphysis pubis to the floor.
6. *Shoulder breadth.* The distance between the most lateral margins of the acromion processes of the scapula.
7. *Waist breadth.* The distance between the lowest lateral points of the costal arch.
8. *Pelvic breadth.* The distance between the two most laterally projecting points on the crests of the ilia.
9. *Maximum head circumference.* From the area between the eyebrows (glabella) around the maximum projection of the occiput.
10. *Neck circumference.* At the level of the thyroid cartilage.
11. *Axillary circumference.* The tape is applied well up on the axillary fossae, and the mean of readings at inspiration and expiration is obtained.
12. *Waist circumference.* The plane where the trunk circumference is minimal is located, and the mean of readings at inspiration and expiration is obtained.
13. *Pelvic circumference.* From the anterior pubic region around the most prominent portion of the buttocks.
14. *Head height.* Calculated by deducting chin height from standing height.
15. *Neck height.* Calculated by deducting suprasternal height from chin height.
16. *Trunk length.* Calculated by deducting symphysis height from suprasternal height.

Various standard growth indices were calculated from the anthropometric measures. Since there was no way of knowing which indices would best reflect changes in configuration, all those used by Zeller and certain

others reported in the literature as being sensitive measures of body build were tentatively employed. The initial battery consisted of 13 indices listed below. The first 12 were expressed as a percentage of standing height to control the factor of height, and the 13th (trunk/extremities) was expressed as a percentage of symphysis height.

- | | |
|-----------------------------|----------------------------------|
| 1. Head height index | 8. Waist breadth index |
| 2. Head circumference index | 9. Pelvic breadth index |
| 3. Neck height index | 10. Axillary circumference index |
| 4. Neck circumference index | 11. Waist circumference index |
| 5. Trunk length index | 12. Pelvic circumference index |
| 6. Leg length index | 13. Trunk/extremities index |
| 7. Shoulder breadth index | |

Results

Somatoscopic findings. The first sorting task required that the judges sort the 69 photographs into early, intermediate, and middle childhood figures. Two analyses were made, the first being a comparison of the judges' ratings with *E*'s rating. The hypothesis under test was that the judges' rating agree with *E*'s rating on one-third of the decisions. Over-all agreement was appreciably higher than one-third ($p < .001$ by binomial test).⁴

The second analysis involved a consideration of the internal consistency of the judges' decisions. Four of the five judges agreed on the specific category of a *S* on 50 per cent of the *S*s, and three or more judges agreed on the classification with dissenting judgments no further removed than the adjoining categories on 83 per cent of the *S*s. Agreement was highest, as might be expected, where the *S*s were at the extremes of the developmental continuum, i.e., early and middle childhood figures. Over-all, it appeared that the percentage of agreement was sufficiently high to enable a further analysis of the elements entering into configuration, and the *S*s three or more judges had agreed upon were isolated for further study. Table 1 is a distribution of these *S*s by sex and age. It may be seen that, while the mean age increases for the *S*s in each developmental category, there is considerable overlapping. These findings support the thesis that configuration is identifiable by inspection, changes lawfully with age, yet is not simply and directly associated with chronological age. Figures 2 to 7 are typical examples of male and female *S*s in each configuration category.

The purpose of the second sorting task was to gain further evidence on the problem of overlap between chronological age and configuration in the developing child. The age factor was therefore controlled by matching *S*s in pairs. The original photographs for the 12 pairs have been placed on

⁴ All probabilities in this paper were calculated by the method of binomial expansion, unless otherwise noted. Probabilities were derived from tables prepared by the National Bureau of Standards (24), except that when the number of items was greater than 49, the method of approximation from the normal curve was used (6).

CHILD DEVELOPMENT

TABLE I

JUDGES' GROUPING OF SUBJECTS FROM PHOTOGRAPHS: AGE RANGES, MEANS, AND STANDARD DEVIATIONS

<i>Configuration</i>	<i>Sex</i>	<i>N</i>	<i>Age Range</i>	<i>Mean Age and SD</i>
Early Childhood	M	6	4- 7 to 6- 2	5- 3 \pm 7 mos.
	F	10	4- 6 to 6- 0	5- 4 \pm 6 mos.
Intermediate	M	9	4- 8 to 6- 3	5- 5 \pm 7 mos.
	F	12	4- 7 to 6-10	5-10 \pm 8 mos.
Middle Childhood	M	11	4- 8 to 7- 2	6- 3 \pm 6 mos.
	F	9	5- 4 to 7- 3	6- 2 \pm 7 mos.

file with ADI (see Figures A1 to L2⁵). Since we wished to test at the same time whether facial or body features provide the more pertinent clues to configurational diagnosis, the judges were confronted with the pictures twice, once sorting them into the more mature and the more immature partners by facial features only and once by body features only. In both sorting tasks the judges' performance proved reliable and consistent ($p < .001$ for both tasks). The sorting by body with face concealed proved to be even more reliable than sorting by face alone, with practically perfect agreement among the judges. These results suggest that children of equal chronological age can be in different maturational stages of body configuration and that their relative maturity can be reliably determined by inspection of either face or body.

When interrogated concerning their choice of the immature face, the judges commented as follows: "by the baby-face"; "by the innocent expression"; "by the tiny peaked jaw"; "the immature child is less filled out around the mouth"; "something about the expression of the eyes, it is not as assured as in the mature ones"; "the face is more rounded, the forehead is more prominent and the chin less so"; "the mature child has more individual expression, the immature one has more of a 'blank' look"; and "in the mature child the roundness is lost, the face takes shape, the mandible and the zygomatic processes show up, and the eye sockets deepen." On the sorting task involving only the body, the judges indicated that they were mainly guided by sloping shoulders, a smooth body surface, and the absence of the waist.

The third problem in terms of sorting concerned the relation of configuration to Ss' sex. It was thought at the outset of this study that the developmental changes in body configuration lead from a rather amorphous infant figure to figures clearly differentiated for the two sexes in middle

⁵ See footnote 2.

MARIA D. SIMON

childhood. In line with this hypothesis it would be expected that the judges would make more errors in classifying immature Ss than in classifying the more mature ones.

The results were contradictory to the hypothesis: the judges could, on the whole, reliably determine the sex of the 69 Ss irrespective of configurational category ($p < .01$ for boys and $p < .001$ for girls with the early childhood figures; $p < .02$ for boys and $p < .001$ for girls with the intermediate figures; and $p < .001$ for boys and $p < .02$ for girls with the middle childhood figures). There was, however, a tendency to mistake certain boys with the early or intermediate figures for girls and several girls with the middle childhood figure for boys. Thus, while the boys apparently grow more masculine as they develop, at least some of the girls seem to grow more masculine as well. In our photographic sample most of the middle childhood types were chronologically very young; and further scrutiny of the photographs revealed that there is apparently a general slimming down in physique at the beginning of the middle childhood period, making boys' and girls' figures similar for a time. Only during the latter stage of this period do girls begin to develop the more typical feminine physical characteristics (see ADI Figures M₁ and M₂^a).

Somatometric findings. The medians and semi-interquartile deviations for the 13 indices are given in Tables 2 and 3. The indices were expressed

TABLE 2
MEDIAN AND SEMI-INTERQUARTILE RANGES FOR ALL INDICES
BY AGE GROUPS: MALES

Indices	A G E G R O U P					
	4-6 to 4-11 (N=8)	5-0 to 5-5 (N=5)	5-6 to 5-11 (N=9)	6-0 to 6-5 (N=8)	6-6 to 6-11 (N=10)	7-0 to 7-5 (N=13)
Head Height . . .	18.2 ± 0.6	18.2 ± 0.82	17.1 ± 4.4	17.0 ± 0.05	16.8 ± 0.49	16.0 ± 0.38
Head Circumf. . .	48.4 ± 1.48	46.0 ± 1.74	46.2 ± 1.1	45.2 ± 0.75	42.9 ± 0.92	42.4 ± 0.52
Neck Height . . .	3.4 ± 0.75	3.4 ± 0.42	3.8 ± 0.30	3.2 ± 0.45	3.6 ± 0.14	4.0 ± 0.24
Neck Circumf. . .	24.1 ± 0.72	23.9 ± 0.62	23.7 ± 0.58	23.2 ± 0.80	22.5 ± 0.96	22.4 ± 0.40
Trunk Length . .	29.2 ± 0.52	28.4 ± 1.22	29.9 ± 0.62	29.9 ± 0.80	29.3 ± 0.75	30.5 ± 0.42
Leg Length	54.8 ± 0.88	53.8 ± 0.38	55.8 ± 0.45	55.6 ± 0.50	56.8 ± 1.2	57.0 ± 0.89
Shoulder Breadth	22.4 ± 0.76	22.8 ± 0.08	23.6 ± 0.58	22.9 ± 0.42	21.8 ± 1.05	21.7 ± 0.28
Waist Breadth . .	15.0 ± 0.25	15.4 ± 0.16	15.7 ± 0.75	14.8 ± 0.50	13.8 ± 0.95	13.9 ± 0.32
Pelvic Breadth . .	16.8 ± 0.35	16.7 ± 0.18	16.6 ± 0.48	16.3 ± 0.40	15.5 ± 0.40	15.2 ± 0.62
Axillary Circumf.	51.1 ± 1.12	51.4 ± 0.56	52.3 ± 1.62	50.5 ± 1.42	49.4 ± 2.35	50.3 ± 1.42
Waist Circumf. . .	46.1 ± 0.98	47.6 ± 0.62	48.4 ± 1.42	46.0 ± 1.62	44.1 ± 2.00	45.8 ± 2.82
Pelvic Circumf. . .	51.9 ± 0.52	53.2 ± 0.52	53.2 ± 1.98	52.4 ± 1.68	50.6 ± 2.01	51.9 ± 1.78
Trunk Extremity	102.5 ± 3.75	102.0 ± 0.01	102.5 ± 2.50	102.8 ± 4.79	99.3 ± 3.40	102.0 ± 1.65

^a See footnote 2.

CHILD DEVELOPMENT

TABLE 3

MEDIANS AND SEMI-INTERQUARTILE RANGES FOR ALL INDICES
BY AGE GROUPS: FEMALES

Indices	A G E G R O U P					
	4-6 to 4-11 (N=5)	5-0 to 5-5 (N=10)	5-6 to 5-11 (N=7)	6-0 to 6-5 (N=8)	6-6 to 6-11 (N=9)	7-0 to 7-5 (N=8)
Head Height . . .	17.6 ± 0.42	17.0 ± 0.50	17.3 ± 0.34	16.7 ± 0.52	16.0 ± 0.88	16.3 ± 0.30
Head Circumf. . .	47.4 ± 1.38	47.6 ± 1.00	46.0 ± 0.62	44.6 ± 1.28	43.0 ± 1.00	43.9 ± 1.12
Neck Height . . .	4.6 ± 0.35	4.2 ± 0.74	3.6 ± 3.25	3.8 ± 0.42	3.7 ± 0.80	4.0 ± 0.23
Neck Circumf. . .	23.3 ± 0.22	22.7 ± 0.49	23.4 ± 0.26	22.3 ± 0.52	22.4 ± 0.50	21.8 ± 0.33
Trunk Length . .	30.1 ± 1.22	29.7 ± 0.94	29.6 ± 0.72	29.6 ± 0.52	29.4 ± 0.52	29.2 ± 0.67
Leg Length	55.0 ± 1.02	55.2 ± 0.70	54.9 ± 0.62	55.6 ± 0.48	56.4 ± 0.05	57.3 ± 0.61
Shoulder Breadth	22.8 ± 1.02	22.9 ± 0.60	23.6 ± 0.50	22.5 ± 0.52	22.5 ± 0.38	22.3 ± 0.68
Waist Breadth . .	15.3 ± 0.88	15.2 ± 0.55	15.5 ± 0.98	14.3 ± 0.39	14.3 ± 0.78	14.6 ± 0.28
Pelvic Breadth . .	17.2 ± 0.22	16.7 ± 0.35	17.5 ± 0.68	16.2 ± 0.45	17.2 ± 0.48	16.5 ± 0.38
Axillary Circumf.	51.7 ± 1.82	50.8 ± 1.90	51.6 ± 1.62	49.2 ± 1.92	51.8 ± 1.28	51.8 ± 1.28
Waist Circumf. . .	47.0 ± 0.72	47.2 ± 0.86	48.0 ± 1.82	44.5 ± 2.02	45.6 ± 1.50	46.5 ± 1.58
Pelvic Circumf. .	53.9 ± 2.08	50.0 ± 0.70	54.6 ± 1.02	51.4 ± 1.88	55.0 ± 2.12	53.9 ± 1.62
Trunk-Extremity	106.5 ± 0.53	104.0 ± 4.54	103.8 ± 0.31	100.0 ± 3.51	101.0 ± 4.35	94.8 ± 1.76

in medians rather than means because most of the distributions of indices were positively or negatively skewed. This skewness seems to be inherent in most anthropometric data, as has been found by other investigators (2, 14).

To test the statistical significance of the differences between the younger and older Ss, the percentage of indices that changed in the expected direction were first computed. The percentage of change in the expected direction was found to depart at a statistically significant level of confidence from the theoretical base rate of 50 per cent ($p < .02$ for boys and $p < .05$ for girls). The "expected direction" was determined on the basis of Zeller's findings and the photographs of this study. It was postulated with increasing maturity that (a) head height, head circumference, neck circumference, trunk length, shoulder breadth, axillary circumference, waist breadth, waist circumference, and the trunk/extremities indices would decrease progressively; (b) the two pelvic indices would decrease in the male Ss, but they would decrease at first and then increase in the female Ss; and (c) the neck length and leg length indices would increase. A word of explanation is in order concerning the directional changes of the shoulder and chest indices. Although the upper thorax and shoulders appear to broaden with age, broad shoulders are dependent for a large extent upon the outlines of the trapezius and deltoid muscles, which do not offer reliable points of

MARIA D. SIMON

TABLE 4

MEDIANS AND SEMI-INTERQUARTILE RANGES FOR ALL INDICES
BY CONFIGURATION TYPE: MALES AND FEMALES

Indices	C O N F I G U R A T I O N					
	EARLY CHILDHOOD		INTERMEDIATE		MIDDLE CHILDHOOD	
	Males (N=6)	Females (N=10)	Males (N=9)	Females (N=12)	Males (N=11)	Females (N=9)
Head Height . . .	17.2 ± 0.8	17.0 ± 0.5	17.8 ± 0.58	17.2 ± 0.52	17.1 ± 0.52	16.7 ± 0.38
Head Circumf. . .	48.4 ± 1.4	46.8 ± 1.21	47.3 ± 0.98	45.4 ± 1.42	45.4 ± 1.58	45.1 ± 1.22
Neck Height . . .	3.5 ± 0.19	4.0 ± 0.6	3.1 ± 0.32	3.8 ± 0.45	3.8 ± 0.6	4.0 ± 0.38
Neck Circumf. . .	24.5 ± 0.6	22.9 ± 0.4	24.0 ± 0.88	22.7 ± 0.55	22.8 ± 1.02	22.1 ± 0.68
Trunk Length . .	29.4 ± 1.05	29.6 ± 0.60	29.3 ± 0.58	29.3 ± 0.65	29.2 ± 0.78	29.6 ± 0.42
Leg Length	55.6 ± 0.65	55.0 ± 0.36	54.9 ± 0.38	55.6 ± 1.12	55.6 ± 0.98	55.8 ± 0.68
Shoulder Breadth	22.1 ± 0.45	22.8 ± 0.75	23.0 ± 0.38	22.8 ± 0.65	22.8 ± 0.78	23.3 ± 0.48
Waist Breadth . .	15.3 ± 0.24	15.2 ± 0.50	14.9 ± 0.18	14.8 ± 0.75	14.5 ± 0.67	14.2 ± 0.41
Pelvic Breadth . .	16.9 ± 0.24	16.9 ± 0.41	16.6 ± 0.52	16.9 ± 1.06	16.1 ± 0.44	16.6 ± 0.31
Axillary Circumf.	51.2 ± 1.20	49.3 ± 1.90	51.4 ± 1.48	51.0 ± 1.90	50.6 ± 1.18	49.5 ± 1.28
Waist Circumf. . .	48.4 ± 1.8	46.7 ± 1.80	46.5 ± 0.68	46.9 ± 1.05	44.4 ± 2.98	44.1 ± 1.98
Pelvic Circumf. .	53.0 ± 0.89	54.2 ± 2.15	52.1 ± 0.62	54.1 ± 2.58	49.4 ± 1.64	53.0 ± 1.72
Trunk-Extremity	101.0 ± 3.0	103.0 ± 3.5	103.0 ± 1.25	99.7 ± 4.15	101.9 ± 3.32	103.0 ± 0.79

measurement. The measure employed in this study was the distance between the extreme lateral margins of the scapulae which presumably would narrow relative to the total height during development.

Table 4 combines the somatoscopic and somatometric approaches and presents the medians and semi-interquartile deviations of the various indices for the three configurational categories derived from judges' ratings. It was found with a statistical analysis similar to that employed for Table 2 and 3 that the median indices for boys changed significantly from early childhood figures to the middle childhood figures ($p < .05$). For the girls the trend of the change in indices was in the expected direction, but it was not statistically significant (eight of the 13 indices changed in the expected direction and there were three ties that were counted against the hypothesis).

Scrutiny of the indices in Tables 2 to 4 (inclusive) reveals that some indices are more useful than others. The most sensitive indices were head circumference (decreasing), waist circumference (decreasing), and leg length (increasing). These findings suggested that an index constructed by dividing head circumference (or waist circumference) by leg length might be a particularly sensitive instrument for registering changes in body configuration. If this is true, one would expect this index to decrease with maturation proportionate to the increase in leg length and the decrease in head circumference (or waist circumference) relative to total stature.

CHILD DEVELOPMENT

Tables 5 and 6 show the median indices for Ss by age groups and configuration groups for the head circumference / leg length index. The same computations were also made for the waist circumference / leg length index (see ADI Tables A and B⁷). Both of these indices tended to decrease with age and advancing configuration. The head circumference / leg length index was somewhat more sensitive to changes in physical maturation than the waist circumference / leg length index.

TABLE 5
HEAD CIRCUMFERENCE / LEG LENGTH INDEX BY AGE GROUPS
(MEDIAN AND SEMI-INTERQUARTILE RANGES)

Age Group	N	Males	N	Females
4-6 to 4-11	8	88.3 ± 2.80	5	85.6 ± 2.92
5-0 to 5-5	5	89.8 ± 3.68	10	85.6 ± 3.08
5-6 to 5-11	8	82.4 ± 2.35	5	82.4 ± 2.02
6-0 to 6-5	8	81.6 ± 1.86	11	80.2 ± 3.12
6-6 to 6-11	10	75.6 ± 3.40	9	78.0 ± 2.58
7-0 to 7-5	3	74.9 ± 1.82	8	75.2 ± 1.40

TABLE 6
HEAD CIRCUMFERENCE / LEG LENGTH INDEX BY CONFIGURATION
(MEDIAN AND SEMI-INTERQUARTILE RANGES)

Configuration	N	Males	N	Females
Early Childhood	6	86.3 ± 4.35	10	85.2 ± 2.76
Intermediate	9	85.2 ± 2.98	12	81.4 ± 3.65
Middle Childhood	11	81.4 ± 4.48	9	80.8 ± 2.08

THE SECOND EXPERIMENT

Procedure

The Ss were 50 highly successful and 50 failing Caucasian students in the first grade, designated as Groups A and B, respectively. The Ss ranged in age from 6-4 to 7-6 and were drawn from five public schools in metropolitan Little Rock, representing all socioeconomic levels. Principals in each school were asked to nominate their 10 "most successful" and 10 "least successful" students on the basis of all-around success or failure, considering academic achievement, social and emotional adjustment, and motivation, excluding IQ. This phase of the study was carried out in the last month

⁷ See footnote 2.

MARIA D. SIMON

of the school year (May, 1958). Each *S*'s IQ on the Pintner-Cunningham Primary Mental Test was obtained from the schools' records. Photographs of these *S*s were not made, and the findings are based on anthropometric measures, identical to those of the first experiment.

Results

Table 7, which is similar to Tables 2 to 4 (inclusive), presents the medians and semi-interquartile deviations for the 13 anthropometric indices. It may be seen that the failing students are relatively more immature than the successful students ($p < .02$ for both sexes).

TABLE 7
INDICES OF TOP (GROUP A) AND FAILING (GROUP B) STUDENTS
(MEDIAN AND SEMI-INTERQUARTILE RANGES)

Indices	G R O U P A		G R O U P B	
	Male 7-0 \pm 2 mos. (N=17)	Female 7-2 \pm 4 mos. (N=33)	Male 6-11 \pm 2 mos. (N=28)	Female 6-11 \pm 1 mo. (N=21)
Head Height	16.3 \pm 0.28	15.7 \pm 0.49	16.3 \pm 0.32	16.4 \pm 0.45
Head Circumference	43.2 \pm 0.74	42.8 \pm 1.36	43.4 \pm 1.00	43.4 \pm 0.93
Neck Height	3.9 \pm 0.46	4.1 \pm 0.46	4.1 \pm 0.48	4.1 \pm 0.48
Neck Circumference	22.5 \pm 0.66	21.4 \pm 0.52	22.8 \pm 0.52	21.5 \pm 0.43
Trunk Length	29.6 \pm 0.43	28.9 \pm 0.53	29.8 \pm 0.75	29.0 \pm 0.99
Leg Length	56.8 \pm 0.42	57.3 \pm 0.54	55.9 \pm 1.12	56.9 \pm 1.23
Shoulder Breadth	21.5 \pm 0.48	21.6 \pm 0.63	21.8 \pm 0.45	21.3 \pm 0.44
Waist Breadth	14.0 \pm 0.24	13.8 \pm 0.51	14.4 \pm 0.40	13.9 \pm 0.74
Pelvic Breadth	15.7 \pm 0.46	16.4 \pm 0.42	16.0 \pm 0.55	15.8 \pm 0.41
Axillary Circumference ..	49.8 \pm 1.06	47.7 \pm 1.23	50.6 \pm 1.50	48.1 \pm 1.82
Waist Circumference	44.8 \pm 1.54	43.9 \pm 1.41	46.1 \pm 1.35	45.2 \pm 1.68
Pelvic Circumference	50.0 \pm 1.02	51.7 \pm 1.5	51.2 \pm 1.25	50.9 \pm 1.12
Trunk-Extremity	99.0 \pm 2.33	95.3 \pm 2.12	101.0 \pm 4.26	98.4 \pm 4.28

Since the failing *S*s tend to be somewhat younger than the successful *S*s, it might be thought that school success depends upon age rather than body configuration. A further obscuring factor is the role of intellectual ability. In order to eliminate these variables, six pairs of boys and eight pairs of girls were matched on the basis of age and IQ. Each *S* of a pair was within three months of age and five IQ points of the other *S*. In the pairs that were not closely matched, the failing *S*s were in most instances older or they had a higher IQ. For each *S* there were 13 anthropometric indices, yielding a population of 182 comparisons for the 14 pairs. It was found that the failing students had a statistically significant fraction of indices associated with immaturity ($p < .01$ for boys and $p < .02$ for girls).

CHILD DEVELOPMENT

Since the head circumference / leg length and the waist circumference / leg length indices are sensitive measures of body configuration, it appeared desirable to do a separate analysis of these indices. For the 14 pairs of matched Ss described above (see ADI Table C⁸), the two types of indices were significantly higher for the failing Ss than for the top Ss (by the *t* test, $p < .01$ for the head circumference / leg length index and $p < .001$ for the waist circumference / leg length index).

These findings were further supported by the trends within the two sex groups. Employing the rank test for paired observations (6), it was found that the head/leg and the waist/leg indices were significantly higher for the failing Ss than for the successful Ss ($p < .01$ for boys and $p < .001$ for girls). Table 8 gives the mean indices for the male and female Ss.

TABLE 8
MEANS AND STANDARD DEVIATIONS OF HEAD CIRCUMFERENCE / LEG LENGTH INDICES AND WAIST CIRCUMFERENCE / LEG LENGTH INDICES FOR THE SUCCESSFUL (GROUP A) AND FOR THE UNSUCCESSFUL (GROUP B) PARTNERS OF MATCHED PAIRS
(N: MALES—6 PAIRS; FEMALES—8 PAIRS)

Scholastic Standing	HEAD CIRCUMFERENCE / LEG LENGTH		WAIST CIRCUMFERENCE / LEG LENGTH	
	Males	Females	Males	Females
Group A	74.6 ± 3.27	73.9 ± 2.56	77.3 ± 3.36	75.4 ± 2.31
Group B	82.5 ± 3.41	77.1 ± 4.83	86.2 ± 5.14	81.1 ± 5.38

Comparing Tables 5 and 8, it may be seen that the mean head/leg index of the successful Ss is approximately the same as that of the average child between the ages of 7 to 7-5, and the mean index for failing Ss corresponds to the values obtained for younger children. The indices in Table 6 for the middle childhood Ss are generally higher than those for the top Ss in the first grade, but it should be remembered that the Ss of this group (selected on the basis of photographs) had a median age approximately 9 months younger than the first grade Ss of the second experiment.

DISCUSSION

The first experiment of the present study, based on a sample of children drawn from the upper socioeconomic levels, corroborated Zeller's (25) earlier findings of a transformation in body configuration between the ages of about 5 to 7. The change is identifiable by inspection of either the faces or bodies alone and by anthropometric measurements. While the change

⁸ See footnote 2.

MARIA D. SIMON

is associated with age, there are significant individual differences in the time of the change.

Boys and girls undergo similar changes in configuration, but there are certain important sex differences independent of the head and genitals. It was found that at the beginning of the middle childhood period there is a general "slimming down" in both girls and boys which tends to make the configurations for the two sex groups similar for a time. Hammond (9) has reported evidence that this slimming down in both sexes begins much earlier. Both boys and girls according to his results exhibit a decrease in body fat beginning at about the age of 2. This decrease continues in girls to about the age of 5 and in boys to about the age of 7. It would appear from the present work that this decrease is not sufficient to mask sex differences until the beginning of the middle childhood period.

The second experiment, based on a representative sampling of children from all socioeconomic levels, corroborated the findings of Hetzer (11) of a positive association between physical maturity and school readiness. It was found that children who fail in the first grade, even though their IQ is in the normal range, tend to be more immature than the top students in the first grade. As in the first study, the anthropometric measures were positively correlated with age, but there remained considerable individual variation within the group independent of age.

If it may be assumed that the findings of this study have a more general validity, a number of past observations of a rather disjointed nature may perhaps be understood as manifestations of the psychophysical transformation of childhood. To consider some of these observations, Boynton and Parsons (2) report that "growth of the various parts of the body is very uneven . . . there is an increase in over 50 per cent in arm span for boys after the age of five-and-one-half, while the increase in head circumference is slightly less than 7 per cent." Lucas and Pryor (13) found from measuring several thousand children that the mean width-length index (identical with the pelvic breadth index of this study) decreases for boys from 1 year of age to adulthood, but for girls the mean index decreases up to 7 years of age, increasing thereafter. Garrison (7) says "the importance of the growth in muscles is reflected in the change in strength and muscle coordination that takes place with age . . . the close relationship between the development of bones and muscles has been observed by different measurements. This relationship is a necessary one, since the muscles attached to the bones by means of tendons hold the different parts of the skeleton in place . . . there is a significant spurt of growth in muscle weight from ages five to six." In the mental field, observers have found that a *S* must have reached the mental age of 6-6 before he can effectively be taught to read (17).

It appears from such cumulative evidence that the period around school entry is something like a halfway house on the road towards maturity. At

CHILD DEVELOPMENT

the termination of the psychophysical transformation of adolescence the individual closely approaches his full physical and mental stature; at the psychophysical transformation between 5 and 7 the child reaches school readiness.

From the present study and other evidence cited, it would appear that it takes an over-all developmental age of body and mind of about 6-6 to attain school readiness in our culture. Obviously, physical maturity is more than skin-deep: it is reflected not only in superficial body features but in the maturational status of the central nervous system which in turn underlies such behavior as readiness to submit to restraints and the application to tasks.

While the samples in the present study are perhaps too small to do more than divulge broad trends, the data offer many interesting leads for future investigation. For example, what is the relationship between constitutional types (10, 12, 21) and developmental configuration? From the materials of our study, it seems that, while developmental configurations appear in their characteristic succession in all children, they are to some extent modified by physical constitution: the constitutionally obese child creates a more infantile impression than is warranted by his over-all developmental status, while the lean, long-legged child may look unduly mature. For example, in our photographic sample of pairs of Ss matched for age and sex, partners L1 of a male pair was unanimously rated by the judges to be more mature than partner L2 (*see* ADI Figures L1 and L2⁹); yet by measurement L1 had a larger waist circumference / leg length index than L2 because of his athletic build.

The foregoing considerations bring us to the question of the comparative merits of the somatoscopic and the somatometric approaches. The case just cited suggests that even a measure as sensitive as the waist/leg index has definite limitations, particularly where extremes of constitution cut across developmental configurations. There may be combinations of measures that might be superior to the ones employed hitherto, or perhaps the method of somatotyping might be adaptable to the purpose of developmental diagnosis.

With a larger and more rigorous study involving repeated measurements to eliminate errors, it might be possible to develop a point scale of physical maturity with cutting points to be used in specifying a stage of development. Such a scale could then be used by teachers and professional workers in assigning a developmental age to a child. This would not only make possible better studies of the interrelationships of biological and psychological factors, but it would also have many practical values: in assessing school readiness, in suggesting the need of a medical examination to evaluate nutritional and endocrine factors, and in assessing readiness for a more rapid advancement in school.

⁹ See footnote 2.

Many of these practical values could be achieved without a refined scale of physical maturity, since it has been shown that configuration is identifiable by inspection. Teachers could be trained to recognize the different figures, particularly the extremes. On the basis of such assessment, the minority of children that has not yet attained the school child physique at age 6 could be given a more thorough psychological examination as a further check on their readiness to enter school.

Another promising line of investigation suggested by the present findings is related to the question of intellectual ability. There is no doubt that intellectual giftedness—although only one of several factors in school readiness—contributes positively towards school success. While all the matched pairs were approximately within one standard deviation of average intelligence, there were no failing children of high IQ to match the top children and there were no top children of low IQ to match the failing children. This is not surprising, since the IQ of a child depends upon two factors inseparably interwoven: the component of maturation and the component of power or capacity. The Binet type scales stress the maturational aspect of a child's intellectual status by defining his score in terms of accelerated or retarded development (mental age); the Wechsler tests define a child's intelligence in terms of "power" by expressing his score as his standing relative to his contemporaries. In any given case, however, we are not at present able to disentangle just how much is contributed by maturation and how much by intellectual giftedness towards a child's IQ. The concept of diagnosis of physical maturity, developed in this paper, may possibly offer a way of separating the maturational and the power components of intelligence. We could perhaps assume that the IQ's of Ss matched for age and physical maturity status would approximately reflect true differences in mental ability.

Such a better understanding of the meaning of an IQ score has not only theoretical but considerable practical interest. Suppose that we have a child of 4 with the typical configuration of early childhood. Suppose further that this child has taught himself to read and write and has an IQ of 150. It might be surmised that this child has not really reached a "mental age" of 6 years, but that he has high intellectual ability compared with his peers, which he will retain throughout life. By contrast, consider another child of the same age and IQ, but whose body configuration approaches the pattern of middle childhood. In this case, the high IQ might reflect an accelerated pattern of growth, and it would be predicted that the child has a relatively high probability of regressing toward the norm as he matures. At least, we would expect the first child to exceed the second child in the performance of mental tasks as the two children mature.

As a last example of an application, the evaluation of physical maturity would seem to be important also in the study of personality. A variety of interesting problems arise: (a) the impact of the school environment on the

CHILD DEVELOPMENT

early and middle child figures; (b) the emergence of personality characteristics in relationship to physical maturity; and (c) better measures of personality differences by equating children for physical maturity, sex, age, and IQ. If personality is conceived as the product of the dynamic interaction of an individual with his culture, then it would appear that the factor of physical maturity would play a significant role.

SUMMARY

This research consisted of two separate but related cross-sectional studies. In the first study, based on 90 Caucasian children from higher socioeconomic levels, it was found that there is a striking change in body configuration between the ages of 4-6 and 7-5. Expert judges discriminated three types of body figures in inspecting photographs of these children, termed the early, intermediate, and middle childhood figures. The identification of the change in configuration was shown not to depend upon age, sex, or particular aspects of the head or body alone. It was also found that judges could reliably differentiate the sex of a child in photographs in which the head and genitals were concealed irrespective of the type of figure. The errors made in sex classification were concentrated on obese boys with the early or intermediate figure and slender girls with the middle figure. Various anthropometric indices also registered the change in body configuration. Of the indices studied, head circumference / leg length and waist circumference / leg length appeared to be most sensitive.

In the second study, based on 50 failing and 50 top Caucasian students in the first grade, drawn from all socioeconomic levels, it was found that failing students tended to be more immature on a battery of anthropometric indices than successful students. This difference did not depend entirely on age or IQ, since failing students matched for age and IQ with the top students were also found to be more immature. Of the measures studied, head circumference / leg length and waist circumference / leg length were the most sensitive indicators of school readiness.

On the basis of these findings, the possibilities for technical refinements in the method of evaluation of physical maturity, and their applications as an aid in the assessment of intelligence and of school readiness, were discussed.

REFERENCES

1. ASHLEY-MONTAGU, M. F. *An introduction to physical anthropology*. Springfield, Ill.: Charles C. Thomas, 1945.
2. BOYNTON, P. L., & PARSONS, R. F. Pupil analysis in the Peabody Demonstration School. *George Peabody Coll. Bull.*, 1933.

MARIA D. SIMON

3. COURTIS, S. A. Maturation as a factor in diagnosis. *Natl Soc. Stud. Educ.*, 1935, 34, 169-174.
4. COURTIS, S. A. What is a growth cycle? *Growth*, 1937, 1(3), 155-174.
5. DEARBORN, W. F., & ROTHNEY, W. M. *Predicting the child's development*. Cambridge, Mass.: Sci-Art, 1941.
6. EDWARDS, A. L. *Statistical methods for the behavioral sciences*. New York: Rinehart, 1954.
7. GARRISON, K. C. *Growth and development*. New York: Longmans-Green, 1952.
8. HAMMOND, W. H. Physique and development of boys from different types of schools. *Brit. J. prev. soc. Med.*, 1953, 7, 231-239.
9. HAMMOND, W. H. Measurement and interpretation of subcutaneous fat with norms for children and young adult males. *Brit. J. prev. soc. Med.*, 1955, 9, 201-211.
10. HAMMOND, W. H. The constancy of physical types as determined by factorial analysis. *Hum. Biol.*, 1957, 29, 40-61.
11. HETZER, HILDEGARD. *Die seelischen Veränderungen des Kindes bei dem ersten Gestaltwandel*. Leipzig: Barth, 1936.
12. KRETSCHMER, E. *Körperbau und Charakter*. Berlin: Springer, 1921.
13. LUCAS, W. P., & PRYOR, H. B. Range and standard deviations of certain physical measurements in healthy children. *J. Pediat.*, 1935, 6, 533-545.
14. MARESH, MARION. Longitudinal evaluation of skeletal maturation. In: *Physical and behavioral growth*. Columbus, Ohio: Ross Laboratories, 1958.
15. MARTIN, R. *Lehrbuch der Anthropologie*. (2nd Ed.) Jena: Fischer, 1942.
16. METTLER, F. A. *Culture and the structural evolution of the neural system*. New York: American Museum of Natural History, 1956.
17. MILLARD, C. V. *Child growth and development in the elementary school*. Boston: Heath, 1951.
18. MORPHETT, MABEL V., & WASHBURN, C. When should children begin to read? *Elem. Sch. J.*, 1931, 31, 496-503.
19. OLSON, W. C., & HUGHES, B. O. Growth of the child as a whole. In: R. G. Barker, J. S. Kounin, & H. F. Wright, (Eds.) *Child behavior and development*, New York: McGraw-Hill, 1943. Pp. 199-208.
20. Report of the special commission on psychodynamic principles. In: *The psychiatrist, his training and development*. Washington: American Psychiatric Association, 1953. Pp. 22-48.

CHILD DEVELOPMENT

21. SHELDON, W. H. *The varieties of human physique*. New York: Harper, 1940.
22. STODDARD, J. D., *et al.* Review of literature on mental and physical development. *Rev. Educ. Rec.*, 1933, 3(2), 1936, 6(1).
23. TERMAN, L. M. Genetic studies of genius, Vol. 1, *Mental and physical traits of a thousand gifted children*. Stanford, Calif.: Stanford Univer. Press, 1925.
24. U. S. Dept. of Commerce. *Tables of the binomial probability distribution*. Washington: U. S. Govt. Printing Office, 1949, (1952 rev.).
25. ZELLER, W. *Der erste Gestaltwandel des Kindes*. Leipzig: Barth, 1936.

PERFORMANCE ON A DIFFICULT DISCRIMINATION FOLLOWING PRETRAINING WITH DISTINCTIVE STIMULI¹

CHARLES C. SPIKER*
State University of Iowa

The results of a previous experiment (4) suggested that some of the difficulty which preschool children have in learning a discrimination problem involving highly similar stimuli may be attributed to factors other than stimulus generalization. One of several plausible hypotheses is that highly similar discriminanda retard the development of appropriate orienting responses. If the stimuli differ with respect to brightness, for example, a small difference may result in the Ss' orienting toward irrelevant features of the stimuli, whereas a large difference may result in a relatively rapid orientation toward the brightness dimension. If this hypothesis is correct, preliminary training with stimuli differing markedly with respect to a given dimension should facilitate subsequent performance with stimuli differing only slightly on this dimension. The present paper reports two experiments concerned with a test of this prediction.

EXPERIMENT I

The first experiment involved a total of three stimuli, a bright, a medium-bright, and a dim light. Group I-C (control) was given 48 trials during which the Ss were rewarded for selecting the bright light when it was presented with the medium-bright light. Group I-E (experimental) was first given 24 trials during which the bright light was positive and the dim one was negative; it was then given 24 trials on the problem which

* Iowa Child Welfare Research Station, State University of Iowa, Iowa City.

¹ The data of the experiments reported in this paper were collected by Hayne W. Reese, to whom the writer is indebted for several valuable suggestions concerning procedures.

CHILD DEVELOPMENT

Group I-C received throughout—the selection of the bright light when presented with the medium one. The performances of the two groups were compared for the final 24 trials when they performed on the same problem.

Method

Subjects. The Ss were 34 children from the preschool laboratories of the Iowa Child Welfare Research Station. On the basis of CA, these Ss were divided into a younger group of 16 Ss and an older second group of 18 Ss. One-half the Ss in each age group was randomly assigned to Group I-E and the other half to Group I-C. The design is thus 2×2 factorial with age and experimental conditions as the main factors. The ranges and means of age for each subgroup of Experiment I are given in the upper half of Table 1.

TABLE 1
THE RANGES AND MEANS OF CA (IN MONTHS) FOR THE SUBGROUPS
OF EXPERIMENTS I AND II

	GROUP E		GROUP C	
	Range	Mean	Range	Mean
<i>Experiment I</i>				
Young	44 to 51	48.5	44 to 52	48.1
Old	53 to 60	56.1	53 to 60	57.3
Combined	44 to 60	52.5	44 to 60	53.0
<i>Experiment II</i>				
Young	45 to 49	47.5	44 to 51	46.9
Old	51 to 63	54.0	52 to 64	56.1
Combined	45 to 63	50.8	44 to 64	51.5

Apparatus. The apparatus was a device for the simultaneous presentation of two lights. It consisted of a box 14 by 14 in. on the front face and 10 in. deep. Two circular apertures, 2 in. in diameter and covered with flashed opal glass, were located on the front face approximately 8 in. apart and $2\frac{1}{2}$ in. from the top. Directly below each of the apertures was a wooden disc, 2 in. in diameter, attached to the leaf of a microswitch. A copper tube, projecting from the front of the box, led to a transparent container into which marbles could be ejected. The controls of the apparatus were placed on the rear of the box and could be operated by E out of view of S.

The electrical circuit of the apparatus was arranged so that E could turn on simultaneously two 75-w. projection lamps, each placed in a light-proof reflecting housing directly behind one of the stimulus apertures. The lamps were operated with variable transformers (Powerstat, Type 10) so that the brightness of each could be independently adjusted for each trial.

CHARLES C. SPIKER

The approximate brightnesses of the bright, medium, and dim lights were 3000, 500, and 15 ft.-candles, respectively. Each of the three brightnesses could be set in either of the two apertures. A two-position switch enabled *E* to select which of *S*'s two buttons would activate the marble dispenser on the next trial. If *S* pushed the button that had been set by *E* to be correct for that trial, a marble was automatically delivered to the container and both lights immediately extinguished. If *S* pushed the incorrect button, the lights remained on until *E* extinguished them, no marble was delivered, and the circuit was interrupted so that correction was not possible.

Procedure. The *E* spent several hours in the preschool groups becoming acquainted with the *S*s. Each *S* was then individually invited to participate in a game to win a toy. The *S* was brought to the experimental room and allowed to select from a variety of dime store toys the one which he wished to win. He was then seated before the apparatus and told that he would have to win enough marbles to buy the chosen toy. Next, the *E* turned on the two lights and pointed to the buttons. The *S* was informed that if he pushed the correct button each time he would receive a marble in the container. He was told that the lights would tell him which button he should push each time. The *E* then turned on both lights with the brighter light in the proper aperture for the first trial. The *S* was then permitted to push one of the two buttons.

Group I-E *S*s received 24 noncorrection trials on which the bright light was always reinforced and the dim light was never reinforced. Without any change in procedure, 24 additional trials were administered on which the bright light was reinforced and the dim light was replaced by the medium light. Control *S*s were given 48 noncorrection trials on the bright and the medium lights, with the bright light always correct. The position of the bright light was randomly determined with the restriction that it appear in one position no more than three successive times and that in successive blocks of six presentations it appear equally often on the left and on the right.

If a given *S* developed a consistent button (position) preference, *E* reminded him that the lights would tell him which button was the correct one to push and that sometimes one button was correct and sometimes the other button was correct.

Results

The percentages of correct responses per block of six trials are graphically represented in Figure 1, where the age subgroups have been collapsed. It can be seen that Group I-E had nearly reached its asymptote of performance within the first block of six trials. Group I-C continued to perform only slightly better than the 50 per cent level throughout the experiment. The upper half of Table 2 presents the mean number of correct responses in the first and second blocks of 24 trials for each of the four subgroups of Experiment I. The performance of Group I-E on the first 24 trials is

CHILD DEVELOPMENT

clearly superior to that of Group I-C, indicating that the stimuli given Group I-E on the first 24 trials were truly more distinctive than those given Group C. The older Ss tended to perform better on the first 24 trials, regardless of stimulus similarity, than did the younger Ss.

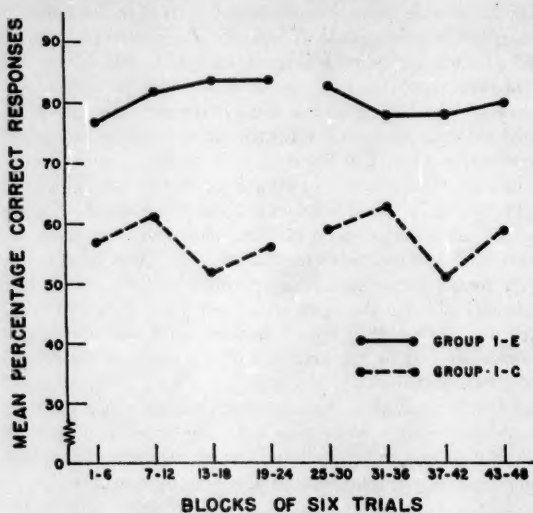


FIGURE 1—Mean percentage of correct responses per block of six trials for the experimental and control groups of Experiment I.

TABLE 2

MEAN NUMBER OF CORRECT RESPONSES IN THE FIRST AND SECOND BLOCKS OF 24 TRIALS

	GROUP E			GROUP C		
	N	First 24	Second 24	N	First 24	Second 24
<i>Experiment I</i>						
Young	8	16.00	15.88	8	10.00	11.75
Old	9	22.11	21.33	9	16.22	15.33
Combined . . .	17	19.24	18.76	17	13.29	13.65
<i>Experiment II</i>						
Young	8	19.88	12.88	8	14.25	10.75
Old	8	19.75	17.00	8	12.50	11.38
Combined . . .	16	19.81	14.94	16	13.38	11.06

CHARLES C. SPIKER

The major hypothesis is concerned with the relative performance of the two groups on the second block of 24 trials. From Table 2 it is apparent that the combined mean for Group I-E is higher than that for Group I-C. Furthermore, the experimental condition appears to have been about equally effective for both the younger and older Ss; that is, there does not appear to be any pronounced interaction between the type of pretraining and the age of the Ss. These trends were evaluated for statistical significance by an analysis of variance, a summary of which is presented in the left half of Table 3. Here it may be seen that the effects of both age and type of pretraining are statistically significant. The interaction of age and type of pretraining does not approach significance.

TABLE 3

ANALYSIS OF VARIANCE OF THE MEAN NUMBER OF CORRECT RESPONSES
IN THE SECOND BLOCK OF 24 TRIALS

Source	EXPERIMENT I				EXPERIMENT II			
	df	Mean Square	F	p	df	Mean Square	F	p
Groups (G)	1	222.62	12.18	.005	1	120.12	8.08	.01
Age (A)	1	173.12	9.47	.005	1	45.12	3.04	
G × A	1	7.45	<1.00		1	24.52	1.65	
Within Groups . . .	30	18.28			28	14.86		
Total	33				31			

The results support the prediction that discrimination learning with highly similar stimuli is facilitated by pretraining in which distinctive stimuli from the same dimension are the relevant stimuli. The facilitation was found even when the performance of Group I-E was compared to that of a control group which received an equivalent amount of practice on the difficult discrimination. The findings are consistent with the notion that pretraining with distinctive stimuli establishes appropriate orienting responses more readily than does pretraining with similar stimuli and that these orienting acts transfer to the more difficult problem. The importance in difficult discrimination of the orienting responses of human Ss has been convincingly demonstrated by Kurtz (1).

EXPERIMENT II

Although Experiment I supports the orienting response interpretation, it should be noted that an alternative interpretation is equally plausible. Group I-E averaged more correct responses during the first 24 trials than did Group I-C, and, conversely, Group I-C averaged more incorrect re-

CHILD DEVELOPMENT

sponses. Thus, the selection of the bright stimulus was more frequently reinforced for Group I-E than for Group I-C, and the selection of the medium-bright stimulus was more frequently nonreinforced for Group I-C than for Group I-E. If the increment in approach tendency resulting from a reinforcement is assumed to be greater than the decrement in approach tendency resulting from a nonreinforcement, the superior performance of Group I-E would be predicted. It is interesting to note that Spence's (2, 3) theory of discrimination behavior assumes that the increment due to reinforcement will be greater than the decrement due to nonreinforcement early in learning when the approach tendencies to the stimuli are relatively weak. Consequently, the Spence theory would predict the major results of Experiment I.

The second experiment was designed to discover whether or not the only function of the distinctive stimulus pretraining, compared to the similar stimulus pretraining, is to produce a relatively greater difference in the strengths of the approach tendencies to the bright and to the medium-bright stimuli. This was accomplished by replicating Experiment I with the single exception that the medium-bright stimulus was made positive for both groups after the first 24 trials. Thus, the experimental group's (Group II-E) greater number of reinforced responses to the bright stimulus during the first 24 trials would result in a relatively stronger tendency to select the incorrect stimulus following the shift. If Group II-E still performs better than the control (Group II-C), the superiority cannot be attributed to a relatively stronger approach tendency to the correct stimulus.

Method

Subjects. The Ss were 32 children from the preschool laboratories of the Iowa Child Welfare Research Station.² On the basis of CA, these Ss were divided into two age groups of 16 Ss each. One-half the Ss in each age group was assigned to Group II-E and the other half to Group II-C. The ranges and means of CA for each subgroup of Experiment II are shown in the lower half of Table 1.

Apparatus and procedures. The apparatus was the same as that used in Experiment I. The procedure for Experiment II differed from that of Experiment I only in that the medium-bright stimulus was made positive for all Ss following the first 24 trials.

Results

The mean percentages of correct responses per block of six trials are graphically presented in Figure 2. Group II-E showed a rapid increase in performance following the first block of six trials. Group II-C, on the other hand, showed a nonsignificant decline in correct responses from the first

² Experiment I was conducted in the spring of 1956; Experiment II, in the spring of 1957. No S included in Experiment I is included in Experiment II.

CHARLES C. SPIKER

to the fourth block of six trials. Both groups lost in performance immediately following the stimulus reversal. Contrary to expectations, Group II-E performed better than Group II-C on the first block of trials following stimulus reversal. This is the result, however, of the relatively more rapid learning of Group II-E during the first six trials as well as throughout the remainder of the task. Group II-C performed slightly better than Group II-E on the first two trials following reversal.

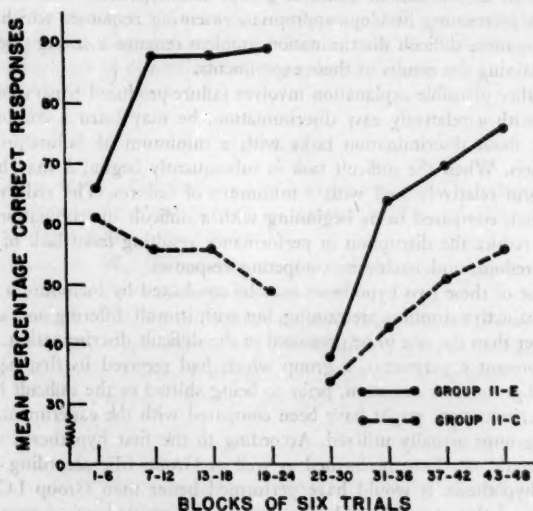


FIGURE 2—Mean percentage of correct responses per block of six trials for the experimental and control groups of Experiment II.

The mean numbers of correct responses in Experiment II are presented in the lower half of Table 2. The performance of Group II-E was superior to that of Group II-C during the first 24 trials, presumably due to the greater similarity of the Group II-C stimuli. The main interest, of course, is in the superior performance of Group II-E during the final 24 trials.

A summary of the analysis of variance, shown in the right half of Table 3, reveals that the performance of Group II-E was significantly better than that of Group II-C during the second block of 24 trials. Neither the main effect of age nor its interaction with type of pretraining was statistically significant, although the former approached significance at the .05 level.

CHILD DEVELOPMENT

DISCUSSION

Considered jointly, the results of Experiment I and Experiment II support the hypothesis that preliminary training with distinctive stimuli facilitates a subsequent discrimination involving similar stimuli differing with respect to the same dimension. The results of Experiment II preclude the explanation of this facilitation solely in terms of the development of a relatively greater difference in approach tendencies for the distinctive stimulus group than for the similar stimulus group. The hypothesis that distinctive stimulus pretraining develops appropriate orienting responses which transfer to the more difficult discrimination problem remains a strong possibility for explaining the results of these experiments.

Another plausible explanation involves failure-produced frustration. If *S* begins with a relatively easy discrimination, he may learn a considerable amount about discrimination tasks with a minimum of failure-produced frustration. When the difficult task is subsequently begun, *S* may be able to perform relatively well with a minimum of failures. The reduction in frustration, compared to *S*s beginning with a difficult discrimination, may in turn reduce the disruption in performance resulting from lack of attention, boredom, and irrelevant competing responses.

A test of these two hypotheses may be conducted by including a group given distinctive stimulus pretraining, but with stimuli differing on a dimension other than the one to be presented in the difficult discrimination. Thus, in the present experiments, a group which had received its first 24 trials on a red-green discrimination, prior to being shifted to the difficult brightness discrimination, might have been compared with the experimental and control groups actually utilized. According to the first hypothesis, such a group would not have performed as well as Group I-E; according to the second hypothesis, it would have performed better than Group I-C. The inclusion of this group would therefore have permitted an assessment of the relative potency of the two hypothesized factors.

SUMMARY

The present paper reports two experiments concerned with the hypothesis that a difficult discrimination learning problem can be made easier by giving the *S*s a preliminary problem involving stimuli differing more markedly with respect to the same dimension.

In Experiment I, 34 preschool children were divided into two age groups. One-half of each age group was assigned to an experimental group (Group I-E) which was given training to respond to a bright light when it was presented with a dim light. The other half of each age group was assigned to a control group (Group I-C) required to respond to the bright light when it was presented with one of medium intensity. All *S*s then received 24 trials with the bright light positive and the medium light negative.

CHARLES C. SPIKER

Group I-E performed significantly better than Group I-C on the final 24 trials.

In Experiment II, 32 preschool Ss were divided into two age groups. One-half of each age group was assigned to Group II-E and the other half to Group II-C. These two groups differed from the corresponding groups of Experiment I only in that the positive stimulus was changed from bright to medium-bright during the final 24 trials. Group II-E performed significantly better than II-C.

Two alternative hypotheses were discussed in accounting for the results: (a) an easy discrimination results in the more effective establishment of appropriate orienting responses, and (b) an easy discrimination prevents the development of disrupting failure-produced frustration.

REFERENCES

1. KURTZ, K. H. Discrimination of complex stimuli: The relationship of training and test stimuli in transfer of discrimination. *J. exp. Psychol.*, 1955, 50, 283-292.
2. SPENCE, K. W. The nature of discrimination learning in animals. *Psychol. Rev.*, 1936, 43, 427-449.
3. SPENCE, K. W. The differential response in animals to stimuli varying within a single dimension. *Psychol. Rev.*, 1937, 44, 430-444.
4. SPIKER, C. C. Effects of stimulus similarity on discrimination learning. *J. exp. Psychol.*, 1956, 51, 393-395.

AMERICAN MEDICAL ASSOCIATION

FOR THE MONTH OF MAY, 1914

VOLUME 41, NUMBER 5

PUBLISHED WEEKLY

CHICAGO, ILL., U.S.A.

Subscription price, \$5.00 per annum in advance

Single copies, 15 cents

Entered as second-class matter, May 2, 1902

Postpaid

Acceptance for mailing at special rate of postage provided for in Act of October 3, 1917

Authorizes sale at wholesale and retail

Copyright, 1914, by American Medical Association

Printed at the Chicago Press, Chicago, Ill.

Second-class postage paid at Chicago, Ill.

Postmaster: This journal is published weekly

except on Sundays and public holidays

and is published at the office of the

editor, 535 North Dearborn Street, Chicago, Ill.

Subscription orders, notices of change of address

and all correspondence should be sent to the

editor, 535 North Dearborn Street, Chicago, Ill.

Entered as second-class matter, May 2, 1902

Postpaid

Acceptance for mailing at special rate of postage provided for in Act of October 3, 1917

Authorizes sale at wholesale and retail

Copyright, 1914, by American Medical Association

Printed at the Chicago Press, Chicago, Ill.

Second-class postage paid at Chicago, Ill.

Postmaster: This journal is published weekly

except on Sundays and public holidays

and is published at the office of the

editor, 535 North Dearborn Street, Chicago, Ill.

Subscription orders, notices of change of address

and all correspondence should be sent to the

editor, 535 North Dearborn Street, Chicago, Ill.

Entered as second-class matter, May 2, 1902

Postpaid

Acceptance for mailing at special rate of postage provided for in Act of October 3, 1917

Authorizes sale at wholesale and retail

Copyright, 1914, by American Medical Association

Printed at the Chicago Press, Chicago, Ill.

Second-class postage paid at Chicago, Ill.

Postmaster: This journal is published weekly

REVERSED SCALES TO CONTROL ACQUIESCENCE RESPONSE SET IN THE PARENTAL ATTITUDE RESEARCH INSTRUMENT

MARVIN ZUCKERMAN*

Indiana University Medical Center

In recent years a number of writers have studied the effect of response sets on objective tests. A response set is a tendency to respond to test items based on something other than the content of the items. The acquiescence response set, or the tendency to use "agree" or "true" categories in a questionnaire, has been shown to play a strong role in tests like the MMPI and the California F scale of Authoritarianism. In a previous study (8), the author of this article studied the influence of three response sets in the Parental Attitude Research Instrument (PARI) (6) and the California F scale (1). These sets were:

1. The Acquiescence Set, or the tendency to agree with oppositely worded items.
2. The Opposition Set, or the tendency to disagree with oppositely worded items.
3. The Extremes Set, or the tendency to use the extreme ratings when offered a choice between "strongly agree" or "strongly disagree" and "mildly agree" or "mildly disagree."

These response sets were shown to have reliability and generality, i.e., a person who agrees with oppositely worded items on the F scale is likely to do so on the PARI as well. Relationships were found between certain response sets and personality traits.

Factor A, "Authoritarian-Control," and Factor B, "Hostility-Rejection," were defined in previous factor analyses of the PARI (7). The Factor A scores are obtained by summing 16 of the 23 individual PARI scales, Factor B scores by summing three scales. The Chapman-Bock method (3) for estimating the influence of acquiescence and content was applied to the data in the previous study (8) and indicated that, although a substantial part of the variance of Factors A and B could be accounted for by acqui-

* Institute of Psychiatric Research, Indiana University Medical Center, Indianapolis.

CHILD DEVELOPMENT

escence set, the content of the items still played the major role in determining response to the items. A subsequent reanalysis of the data using factor analysis (9) indicated greater influence of acquiescence in Factor A and also indicated that the extremes set accounts for part of the variance of this factor. The probable reasons for the quantitative disagreement in these two methods of analysis are given in the latter article. However, in both methods of analysis the acquiescence set was found to play a significant role in the PARI. Since it is important for an investigator to know whether differences he obtains between criterion groups are due to item content or response set, it is desirable to develop a form of the test where the influence of response sets can be controlled and measured. This article will suggest the use of the reversed items developed in the previous study (8) to accomplish the control and measurement of response set in the PARI.

Bell (2) has attempted to accomplish this purpose using a statistical technique. This technique consists of converting a subject's raw scores on the 23 PARI scales to standard scores based on his own mean and standard deviation. Bell interprets the mean on the 23 scales as a measure of acquiescence. In doing so, he overlooks the fact that acquiescence may have different weights in each of the individual scales. Furthermore, the general elevation of the profile is probably due to content as well as acquiescence because 20 of the 23 scales are scored in the same direction and load highly on two group factors. This paper will offer an alternate technique for measuring response to content and the response sets themselves by using the reversed items developed in the previous study (8) along with the original PARI items. The repeat reliabilities of the original and reversed PARI scales and their correlations with each other will be reported. The purpose is to evolve a form of the PARI composed of scales which adequately sample the main factors of the test and correlate reasonably high with their reversed versions.

METHOD AND PROCEDURE

Subjects

The subjects were 88 unmarried student nurses. Their average age was 19.7 ($SD \pm 1.0$). They were above average in intelligence as indicated by a mean A.C.E. score of 61.1 ($SD \pm 23.8$). They were given the usual and reversed form of the PARI prior to receiving other personality tests. Two weeks later 49 of the group, who happened to be in a class that was available, received the two forms of the PARI again in order to obtain repeat reliabilities on the usual and reversed scales.

Tests

1. The usual 23-scale form of the PARI.
2. A reversed form of the PARI developed by this author. The items of 20 of the PARI scales were rewritten in an attempt to reverse their

MARVIN ZUCKERMAN

meanings (see Appendix). The remaining three scales consist of items worded in a positive sense and reflecting democratic-equalitarian attitudes. Most women agree with these items. They were included in the original test because some women complained that there were no items on the test with which they could agree. Since the reversed items of the other scales accomplish this purpose, the author did not consider it worthwhile to develop reversals of these scales.

Christie (4) has pointed out some of the difficulties in reversing the meaning of an item of the F scale. In reversing the items of the PARI, the author recognizes that a low correlation between a scale and its reversed version may indicate either the influence of acquiescence response set or inadequate reversal. However, a high correlation is empirical proof of the adequacy of reversal. The purpose of this study is mainly to select scales which demonstrate reversal in reasonably high correlations with their reversed versions.

RESULTS AND DISCUSSION

Table 1 lists the repeat reliabilities of the usual and reversed PARI scales. The correlations between the pairs of the scales on the original 88 subjects are given with the correlations corrected for attenuation. Since both kinds of scales are scored in the acquiescence direction, a high negative relationship would be expected if the subjects were responding consistently to content, i.e., a subject who scores high on the usual PARI scale should score low on its reversed form. A lower negative correlation or a positive correlation indicates either inadequate reversal of scale and/or the influence of acquiescence response set. It can be seen that the correlations between the pairs of scales are all negative except for scale 19. Furthermore, all of these correlations are significant with the exception of scales 6, 15, and 19. When the correlations are averaged, using the π' conversion, the average correlation between positive and reversed scales is $-.36$. The average of the correlations corrected for attenuation is $-.62$. Chapman and Bock (3) list eight correlations between positive and reversed versions of the F scale. The average correlation of these studies was $-.10$ (the signs were reversed to make the results comparable since in these F scale studies scales were scored in the authoritarian direction while in this study they were scored in the acquiescence direction). Christie (4) criticized these studies because of the manner in which reversed items were worded. He developed his own reversed scales. The average correlation in the eight studies he reports was $-.21$. He presents the correlations corrected for attenuation, and the average of these was $-.39$. It can be concluded that the scales in this study were, on the average, more successfully reversed than the F scale has been to date.

Since the use of reversed scales along with the usual PARI scales would entail doubling the length of the test, it may be desirable to shorten the

CHILD DEVELOPMENT

length by reducing the number of scales. In suggesting a shortened form of the PARI to use with reversed scales, this author had two considerations: (a) the scales used should constitute an adequate sample of the main factors in the test; (b) the scales used should have a relatively high correlation with their reversed scales. The first consideration is complicated by the fact that five factors were extracted in an analysis of PARIs given to unmarried student nurses by Schaefer and Bell (5), while only three factors were extracted from PARIs given to mothers by Zuckerman (7). The suggested version of the PARI contained in Table 2 is a compromise between these two analyses. Factors A and B represent two of the three factors found in mothers. The scales representing the third factor were not re-

TABLE I
REPEAT RELIABILITIES OF USUAL AND REVERSED PARI SCALES AND
THE CORRELATIONS BETWEEN THEM

Scale	TEST-RETEST RELIABILITY		CORRELATION BETWEEN TESTS	
	Usual Scale* (P+)	Reversed Scale* (P-)	P+, P- P+, P-†	(Corrected for Attenuation)
1. Encouraging Verbalization78
2. Fostering Dependency70	.66	-.43	-.63
3. Seclusion of the Mother72	.47	-.30	-.52
4. Breaking the Will80	.68	-.28	-.38
5. Martyrdom64	.50	-.46	-.81
6. Fear of Harming the Baby71	.59	-.15	-.23
7. Marital Conflict71	.61	-.46	-.70
8. Strictness80	.71	-.47	-.62
9. Irritability65	.64	-.53	-.82
10. Excluding Outside Influences70	.72	-.30	-.42
11. Deification63	.69	-.34	-.52
12. Suppression of Aggression73	.78	-.29	-.38
13. Rejection of Homemaking Role73	.55	-.28	-.44
14. Equalitarianism47
15. Approval of Activity74	.56	-.11	-.17
16. Avoidance of Communication62	.61	-.49	-.80
17. Inconsiderateness of Husband74	.68	-.35	-.49
18. Suppression of Sex67	.71	-.45	-.65
19. Ascendancy of the Mother68	.64	.05	.08
20. Intrusiveness52	.64	-.59	-1.02
21. Comradeship and Sharing70
22. Acceleration of Development60	.68	-.37	-.58
23. Dependency of the Mother66	.65	-.36	-.55

* N = 49.

† N = 88.

MARVIN ZUCKERMAN

TABLE 2

SUGGESTED SELECTION OF SCALES TO USE IN CONJUNCTION
WITH REVERSED SCALES

Factors	Subfactors	Scales	LOADING ON THE RELEVANT FACTORS		$r_{P+.P-}$ Corrected for Attenuation
			Zuckerman et al. (7)	Schaefer & Bell (5)	
Factor A. Authoritarian-Control (Zuckerman)					
1. Harsh Punitive-Control (S & B)					
	4	Breaking the Will77	.53	— .38
	10	Excluding Outside Influences80	.55	— .42
	11	Deification74	.45	— .52
2. Suppression and Interpersonal Distance (S & B)					
	16	Avoidance of Communication70	.30	— .80
	18	Suppression of Sex78	.42	— .65
3. Over-possessiveness (S & B)					
	2	Fostering Dependency71	.53	— .63
	5	Martyrdom77	.59	— .81
	20	Intrusiveness73	.52	— 1.02
4. Excessive Demand for Striving (S & B)					
	8	Strictness54	.58	— .62
	22	Acceleration of Development74	.51	— .58
Factor B. Hostility-Rejection (Zuckerman and S & B)					
	7	Marital Conflict34	.47	— .70
	9	Irritability69	.60	— .83
	13	Rejection of the Homemaking Role ..	.62	.64	— .44

Suggested Summary Scores:

- | | | |
|------------------------|----------------|------------------|
| 1. A total (10 scales) | 7. A — total† | 13. A + — total‡ |
| 2. A1 total (3 scales) | 8. A1 — total | 14. A1 + — total |
| 3. A2 total (2 scales) | 9. A2 — total | 15. A2 + — total |
| 4. A3 total (3 scales) | 10. A3 — total | 16. A3 + — total |
| 5. A4 total (2 scales) | 11. A4 — total | 17. A4 + — total |
| 6. B total (3 scales) | 12. B — total | 18. B + — total |

19. **Acquiescence Set:** The number of item pairs where a subject agrees with both the positive and reversed forms of the same item.

20. **Opposition Set:** The number of item pairs where a subject disagrees with both the positive and reversed forms of the same item.

21. **Extremes Set Ratio:**
$$\frac{\text{No. of strongly agree} + \text{No. of strongly disagree responses}}{\text{Total No. of responses on all scales, usual and reversed.}}$$

* Summary Scores are obtained by summing the individual scale scores defining the factor. Scales are weighted equally except for Factor B where scale 7 receives only $\frac{1}{2}$ weight.

† Scores 7-12 are the summary scores on the reversed scales scored in the "agree" direction.

‡ Usual and reversed scores may be combined by subtracting the scores on the reversed factors from the scores on the positive factors and adding a constant large enough to eliminate all negative scores.

CHILD DEVELOPMENT

versed in this study. The four subfactors listed under Factor A together with Factor B represent the five factors found in student nurses. The 13 scales listed sample all of these factors and correlate reasonably high with their reversed scales. The use of the usual and reversed scales together would make the total test consist of 26 scales or 130 items. The usual PARI consists of 23 scales or 115 items.

Suggested ways of scoring this revised PARI are given in Table 2. Major factor, subfactor, or individual scale scores may be used. One can consider the positive and reversed scales separately or combine them to form a score controlled for acquiescence response set. Measures of acquiescence, opposition, and extremes sets may be obtained in the manner indicated.

No evidence is available on the validity of the reversed PARI scales. However, this version of the PARI will enable an investigator to determine whether any results he obtains are due to response sets or item content. It is possible that the response set scores may prove useful in themselves.

SUMMARY

The purpose of this study was to develop a form of the Parental Attitude Research Instrument, controlled for acquiescence response set. The items comprising 20 of the original 23 PARI scales were reversed in meaning, and these reversed scales were correlated with the original scales. Repeat reliabilities were obtained on both types of scale. The correlations between usual and reversed scales indicated fairly good reversals for a number of the scales. A revised form of the PARI was suggested which would include 13 scales from the original PARI, sampling the main factor dimensions of the test and their 13 reversed scales. This version of the test would yield three response set scores in addition to parental attitude scores controlled for acquiescence response set.

APPENDIX

Reversed Scale Items for the Parental Attitude Research Instrument^{1,2}

		Scale No.
(2)	1. A good mother lets her child learn the hard way about life.	2
(3)	2. A good mother should develop interests outside the home.	3
(4)	3. A child should never be taught to fear adults.	4
(5)	4. Parents shouldn't feel they have to sacrifice for their children.	5
(6)	5. Mothers worry too much about bathing babies.	6
(7)	6. There is no reason for arguments in a happy marriage.	7
(8)	7. Strict training will make a child resent his parents later on.	8

¹ These items are administered after the regular PARI. The form for responding and the instructions are the same as in the regular PARI. Scoring is also the same with a "strongly agree" response weighted 4; "mildly agree," 3; "mildly disagree," 2; and "strongly disagree," 1. The number in parentheses preceding each item is the number

MARVIN ZUCKERMAN

		Scale No.
(9)	8. There is no reason why a day with the children should be up- setting.	9
(10)	9. Children have every right to question their mothers' views.	10
(11)	10. Loyalty on the part of children to their parents is something that the parents should earn.	11
(12)	11. Children should be taught to fight so that they can take care of themselves.	12
(13)	12. Taking care of a home doesn't have to coop a woman up.	13
(15)	13. A child needs time just to sit around and do nothing if he feels like it.	15
(16)	14. Children should be encouraged to talk about their problems.	16
(17)	15. Fathers generally are kind and helpful.	17
(18)	16. Children should be taught about sex as soon as possible.	18
(19)	17. It is not the mother's place to make the rules for the home.	19
(20)	18. A child's thoughts and ideas are his own business.	20
(22)	19. Very few children are toilet trained by 15 months of age.	22
(23)	20. A young mother doesn't need any help when going through her first experience with a baby.	23
(25)	21. A child should learn that he has to be disappointed sometimes.	2
(26)	22. A good mother has an active social life.	3
(27)	23. You can't make a child behave by cracking down on him.	4
(28)	24. There is no reason why a mother can't be happy and make her child happy too.	5
(29)	25. Most young mothers don't worry much about handling or hold- ing the baby.	6
(30)	26. A good wife never has to argue with her husband.	7
(31)	27. Strict discipline makes children grow up to be mean or re- bellious.	8
(32)	28. Most mothers never get to the point where they can't stand their children.	9
(33)	29. If a parent is wrong he should admit it to his child.	10
(34)	30. A child should be taught that there are many other people he will love and respect as much or more than his own parents.	11
(35)	31. A child should be taught to fight his own battles.	12
(36)	32. Most mothers are content to be with children all the time.	13
(38)	33. There is plenty of time for children to strive for success after they are older.	15
(39)	34. A child should always be encouraged to talk about his troubles.	16
(40)	35. Husbands have a perfect right to their own interests.	17
(41)	36. There is nothing wrong with bathing boys and girls in the same bath tub.	18
(42)	37. The family is better off when the husband settles most of the family problems.	19

of the item in the PARI form IV [see PARI manual or Appendix in (6)] which is re-versed in the item given here. The second number is the item number for the reversed PARI form. Following each item is the scale number of the scale on which the item is scored. The scale numbers are the same for the usual scales and their reversed versions.

²Permission is granted to anyone to reproduce this material with proper acknowledgement, without permission of the author or of the Society for Research in Child Development.

CHILD DEVELOPMENT

		Scale No.	
(43)	38. Children are entitled to keep their own secrets.	20	(8)
(45)	39. A child should take all the time he wants to before he walks.	22	(8)
(46)	40. Women should handle most of child-raising without much help from others.	23	(8)
(48)	41. Children should be encouraged to undertake tough jobs if they want to.	2	(8)
(49)	42. A mother can keep a nice home and still have plenty of time left over to visit with neighbors and friends.	3	(8)
(50)	43. There is no need for children to look on parents as their bosses.	4	(8)
(51)	44. Most children are grateful to their parents.	5	(8)
(52)	45. Little accidents are bound to happen when caring for young babies.	6	(8)
(53)	46. If a couple really loves each other there are very few arguments in their married life.	7	(8)
(54)	47. If children are given too many rules they will grow up to be unhappy adults.	8	(8)
(55)	48. Most mothers can spend all day with the children and remain calm and even-tempered.	9	(8)
(56)	49. A child should be encouraged to look for answers to his questions from other people even if the answers contradict his parents.	10	(8)
(57)	50. Most children soon learn that their parents were mistaken in many of their ideas.	11	(8)
(58)	51. It's quite natural for children to hit one another.	12	(8)
(59)	52. Most young mothers don't mind spending most of their time at home.	13	(8)
(61)	53. Children should have lots of time to loaf and play.	15	(8)
(62)	54. A mother should be concerned with any problem of a child no matter how trivial.	16	(8)
(63)	55. In most cases the mother rather than the father is responsible for trouble in the home.	17	(8)
(64)	56. Sex play is a normal thing in children.	18	(10)
(65)	57. A mother should take a back seat to her husband as far as the planning is concerned.	19	(10)
(66)	58. A good parent doesn't try to pry into the child's thoughts.	20	(10)
(68)	59. A child needs to be emotionally close to its parents for a long time.	22	(10)
(69)	60. A woman should be on her own after having a baby.	23	(10)
(71)	61. Children have to face difficult situations on their own.	2	(10)
(72)	62. Mothers should get out of the home fairly often.	3	(10)
(73)	63. If a child acts mean he needs understanding rather than punishment.	4	(10)
(74)	64. Children don't "owe" their mothers anything.	5	(10)
(75)	65. Most mothers are confident when handling their babies.	6	(10)
(76)	66. Almost any problem can be settled by quietly talking it over.	7	(10)
(77)	67. Most children are disciplined too much.	8	(10)
(78)	68. Raising children is an easy job.	9	(10)
(79)	69. When a child thinks his parent is wrong he should say so.	10	(10)
(80)	70. A parent should not expect to be more highly esteemed than other worthy adults in their children's eyes.	11	(10)

MARVIN ZUCKERMAN

		<i>Scale No.</i>
(81)	71. Children should be taught ways of defending themselves in a fight.	12
(82)	72. If you run your home right, you have plenty of time to do the things you like to do.	13
(84)	73. It isn't good for children to be constantly running from one activity to another.	15
(85)	74. A mother should always be concerned about upset feelings in a child.	16
(86)	75. Most husbands show good understanding for a mother's problems.	17
(87)	76. Sex is no great problem for children if the parent doesn't make it one.	18
(88)	77. It's up to the father to take charge of the family.	19
(89)	78. Being a mother doesn't give women the right to know everything in their children's lives.	20
(91)	79. Toilet training should be put off until the child indicates he is ready.	22
(92)	80. A woman should be up and around a short time after giving birth.	23
(94)	81. Children should be encouraged to undertake all kinds of jobs no matter how hard.	2
(95)	82. It is important for a mother to have a social life outside of the family.	3
(96)	83. Children have a right to rebel and be stubborn sometimes.	4
(97)	84. Having children doesn't mean you can't have as much fun as you usually do.	5
(98)	85. Mothers shouldn't worry much about calamities that might happen to their children.	6
(99)	86. Husbands and wives who have different views can still get along without arguments.	7
(100)	87. Strict training makes children unhappy.	8
(101)	88. A mother should keep control of her temper even when children are demanding.	9
(102)	89. A good mother can tolerate criticism of herself, even when the children are around.	10
(103)	90. Loyalty to parents is an over-emphasized virtue.	11
(104)	91. Most parents prefer a "scrappy" child to a quiet one.	12
(105)	92. Most young mothers are pretty content with home life.	13
(107)	93. A child should have time to just dawdle or daydream.	15
(108)	94. Anything a child wants to tell a parent is important enough to listen to.	16
(109)	95. Most men try to take their wives out as often as they can.	17
(110)	96. Children are normally curious about sex.	18
(111)	97. Most wives think it best that the husband take the lead in family affairs.	19
(112)	98. Every child should have an inner life which is only his business.	20
(114)	99. The longer a child is bottle or beast fed the more secure he will feel.	22
(115)	100. Any woman should be capable of taking care of a baby by herself.	23

CHILD DEVELOPMENT

REFERENCES

1. ADORNO, T. W., FRENKEL-BRUNSWICK, ELSE, LEVINSON, D., & SANFORD, R. N. *The authoritarian personality*. New York: Harper, 1950.
2. BELL, R. Q. Maternal types identifiable through attitude measures. *Amer. Psychologist*, 1957, 12, 398-399. (Abstract)
3. CHAPMAN, L. J., & BOCK, R. D. Components of variance due to acquiescence and content in the F scale measure of authoritarianism. *Psychol. Bull.*, 1958, 55, 328-333.
4. CHRISTIE, R., HAVEL, JOAN, & SEIDENBERG, B. Is the F scale irreversible? *J. abnorm. soc. Psychol.*, 1958, 56, 143-159.
5. SCHAEFER, E. S., & BELL, R. Q. Patterns of attitudes toward child rearing and the family. *J. abnorm. soc. Psychol.*, 1957, 54, 391-395.
6. SCHAEFER, E. S., & BELL, R. Q. Development of the Parental Attitude Research Instrument. *Child Developm.*, 1958, 29, 339-361.
7. ZUCKERMAN, M., BARRETT-RIBBACK, BEATRICE, MONASHKIN, I., & NORTON, J. Normative data and factor analysis on the Parental Attitude Research Instrument. *J. consult. Psychol.*, 1958, 22, 165-171.
8. ZUCKERMAN, M., & NORTON, J. Acquiescence and extremes sets and their role in tests of authoritarianism and parental attitudes. *Psychiat. Res. Rep.*, 1958, 10, 28-45.
9. ZUCKERMAN, M., & NORTON, J. Response set and content factors in the California F scale and the Parental Attitude Research Instrument. *J. soc. Psychol.*, in press.

DEPENDENCE AND INDEPENDENCE IN THE CHILDREN OF WORKING MOTHERS¹

ALBERTA ENGVALL SIEGEL*

The Pennsylvania State University

LOIS MEEK STOLZ

Stanford University

ETHEL ALICE HITCHCOCK

Illinois State Normal University

and JEAN ADAMSON

Creche and Kindergarten Association of Queensland, Australia

The number of mothers of young children in gainful employment outside the home has risen dramatically in the United States in the last two decades (12). Despite widespread popular and professional concern with the possible effects a mother's working status may have upon her children's personality development, and despite the willingness of many people in mental health professions to make strongly worded statements on this issue, there has been notably little research concerning the children of working mothers (10). Such research data as have appeared have typically not been capable of clear interpretation because in some studies no control groups were used, in others comparisons between the children of working and non-working mothers have been confounded by group differences in such vari-

* Department of Child Development, Pennsylvania State University, University Park.

¹ This study was conducted at Stanford University during 1957-1958 under a grant from the Elizabeth McCormick Memorial Fund to the Committee on Maternal Employment organized by Miss Christine M. Heinig of the American Association of University Women. The authors wish to thank the other members of the Committee—Christine M. Heinig, Margaret McFarland, Leon Yarrow, and Marian Radke Yarrow—as well as Donald Brieland, for suggestions concerning this study. For cooperation in the conduct of the study, we are indebted to the administration, faculty, and parents of the Palo Alto and Sunnyvale School Districts in California. We are especially grateful to Jack Rand, Assistant Superintendent, and John Caffrey, Director of Research, of the Palo Alto School District, and to John Holtorf, Superintendent, of the Sunnyvale School District. Frances R. Shafel of Stanford University gave administrative cooperation to this study, and Eleanor Walker Willemssen gave generous assistance in data processing and analysis.

CHILD DEVELOPMENT

ables as socioeconomic status, family intactness, family size, and in some the status of working and nonworking mothers was vaguely or arbitrarily defined (6, 10).

The purpose of the present study, conceived as an exploratory investigation, was to identify possible differences between the children of working and nonworking mothers with respect to dependence and independence, and to do so within a research design which permits us to isolate children's characteristics associated with the employment status of their mothers from characteristics associated with other features of their lives and only incidentally associated with maternal employment status.

Various manifestations of dependence and independence in young children were the focus of our study because existing theory and evidence concerning socialization suggest that a child's development with respect to dependency is intimately associated with his relations with his parents and especially with his mother. It seemed reasonable to believe that if maternal employment is indeed a significant factor in the constellation of psychological and social factors which provide a background for the mother-child relation and thus for personality development in the young child, its implications will be greatest for the child's development with respect to dependence and independence. Development in this respect may be presumed to relate to the consistency of the care the child receives, the frequency of his contacts with the caretaker, the number of different people who assume caretaking responsibility for him, the diversity of their child-rearing techniques, and the diversity of their attitudes toward the child. All of these conditions may be quite different for the child of a working mother than for the child of a fulltime homemaker.

CORRELATES OF MATERNAL EMPLOYMENT STATUS

To secure information needed for the selection of children to be subjects in the observational study, brief questionnaires were circulated by mail to the mothers of kindergarten children in 17 schools in two large suburban communities. Covered by a letter from the school principal soliciting cooperation, these one-page forms were mailed to 1067 mothers. They were told, "We are cooperating with a research group from Stanford University in collecting data for a study of kindergarten children in relation to adults and other children," and they were asked to use the stamped envelope enclosed in returning the completed questionnaire to the school. The research interest in maternal employment was not mentioned, and certain masking items were included in the questionnaire to attenuate its emphasis on maternal employment. A total of 917 questionnaires were completed and returned, representing an 86 per cent return. From the questionnaire responses it was possible to determine the relation between a woman's employment status and certain other features of her family situation. These survey findings will be reported summarily here.

One of the communities in this study has a large population of college professors and other professionals, but also has the range and diversity of occupational groups typical of many large and well-to-do suburbs in the United States. The other community, also a suburb, has a greater concentration of skilled industrial workers as well as many personnel from a nearby military installation. Both towns are undergoing rapid growth and have many recent arrivals in their populace. A pattern of ethnic segregation by residential area prevails in both communities, and the schools which participated in the study are so located that they drew their pupils overwhelmingly from nonminority families. No ethnic identification of questionnaire respondents was attempted, but our information is that the families circularized included only a sprinkling of minority people, including Negroes and persons of Japanese and Mexican background.

Extent of Maternal Employment

At the time of the study, 11 per cent of the mothers of the kindergarten children were employed full-time, i.e., 32 or more hours per week. An additional 5 per cent were employed part-time, 2 per cent working 16 to 31 hours per week and 3 per cent working 1 to 15 hours per week. The largest single group, 84 per cent of the total number of respondents, were full-time homemakers. These data suggest that, in extent, maternal employment in these communities corresponds closely to the national average (12, p. 69).

The 149 employed mothers were in a wide variety of occupations, ranging from professions (medicine, chemistry, architecture) to relatively unskilled occupations (domestic, laundry, and cannery work). Concentrations of working mothers were observed in teaching, in nursing, in secretarial and clerical work, and in restaurant and factory work.

The total number of months each mother had worked (either full-time or part-time) during the lifetime of her kindergarten-age child was determined. We attempted to classify the employment histories of the mothers, but these histories were so diverse and often so irregular that they defied classification in any reasonable number of groups. Given below are the percentages of respondents whose total number of months of work during the kindergarten-age child's lifetime fell within each of the intervals indicated:

- 64% had never worked during child's lifetime;
- 10% had worked from 1 to 6 months during child's lifetime;
- 7% had worked from 7 to 12 months during child's lifetime;
- 8% had worked from 13 to 24 months during child's lifetime;
- 5% had worked from 25 to 36 months during child's lifetime;
- 3% had worked from 37 to 48 months during child's lifetime;
- 3% had worked from 49 to 60 months during child's lifetime.

It is evident that but a small minority of these mothers (no more than 3 per cent) had been employed throughout the lifetimes of their kindergarten children, and also that some of the mothers who were full-time homemakers

CHILD DEVELOPMENT

at the time of the survey had worked at some earlier period within the lifetime of the kindergarten child.

Maternal Employment and Family Size

In the families surveyed, the median number of children was three. Working mothers were found significantly more often in families with fewer than the median number of children than in families with the median number or more ($\chi^2 = 8.65$, $df = 1$, $p < .01$). However, among the families of working mothers there was no relation between family size and the number of hours the mother worked per week.

Maternal Employment and Ages of Children in Family

The majority of the families surveyed had preschool-age children in the home: 566 families (62 per cent of the respondents) had at least one child younger than the kindergarten child. Working mothers were found significantly more often in families having no child of preschool age than in families including a preschool-age child ($\chi^2 = 43.97$, $df = 1$, $p < .001$). However, among the families of working mothers, there was no relation between age composition of family and the number of hours per week the mother worked.

Maternal Employment and Family Intactness

Of the respondents, 95 per cent represented intact homes (homes with united parents), 4 per cent represented homes broken by divorce or separation, and 1 per cent represented homes broken by death of the father.

The proportion of working mothers was higher in nonintact families than in united families ($\chi^2 = 65.78$, $df = 1$, $p < .001$). This finding does not mean that working mothers were typically from nonintact families, for in fact the great majority of working mothers in the group studied (119 of the 149) were from intact families. Rather, it means that mothers from nonintact families were typically working mothers—of the 49 mothers from nonintact families in our survey, 30 were employed. Interestingly, however, again the number of hours spent at work was not discriminating: working mothers from intact families did not differ in the number of hours they worked per week from working mothers from nonintact families.

It may be appropriate to mention here that family intactness and age composition of family were associated, with proportionally more intact families among those with preschool children ($\chi^2 = 17.80$, $df = 1$, $p < .001$). In addition, family intactness and family size were associated, with proportionally more intact families among those with the median number of children or more than among those with fewer than the median number of children ($\chi^2 = 9.86$, $df = 2$, $p < .01$).

Maternal Employment and Social Level of Husband's Occupation

Several items in the questionnaire elicited information concerning the occupation and type of work of the child's father. The socioeconomic level

of the father's work was rated on the seven-point scale of occupations developed by Hollingshead (4). Such ratings were attempted only for fathers in the 868 intact families concerning whose work the information supplied by the mother was deemed sufficient by two judges rating social level. For 23 intact families the information was judged insufficient, so ratings were made on 845 cases in all.²

The fathers from intact families whose occupational levels could be rated were found to be distributed as follows:

Level 1 (higher executives of large concerns, proprietors of businesses valued over \$100,000, and major professionals)	17%
Level 2 (business managers in large concerns, proprietors of medium-sized businesses, and lesser professionals)	15%
Level 3 (administrative personnel, owners of small businesses, and minor professionals)	16%
Level 4 (clerical and sales workers, technicians, and owners of businesses valued under \$6000)	14%
Level 5 (skilled manual employees, small farmers)	23%
Level 6 (machine operators, semiskilled employees, and tenant farmers)	12%
Level 7 (unskilled employees and sharecroppers)	3%

To determine whether maternal employment was associated with occupational level of father, we compared the frequency distribution of occupational levels of husbands of working mothers with the frequency distribution of occupational levels of husbands of nonworking mothers, using the Kolmogorov-Smirnov test for two independent samples (7, pp. 127-136). No significant differences were observed between the two distributions; in the present survey sample, social level of father's occupation and working status of mother were not associated.

It is of some interest to note that there was also no significant association between family size and social level of father's occupation. That is, among intact families, those above the median in social status (as indicated by social position of father's occupation) did not differ in size from families at or below the median in social status.

COMPARISON OF CHILDREN OF WORKING AND NONWORKING MOTHERS

On the basis of the information contained in the survey questionnaires, matched pairs of children of working and nonworking mothers were

² Working independently, two judges first rated occupational level for 164 cases chosen from the total group. Both indicated their "confidence" in each rating. For those ratings of which Judge A was confident, the correlation between the two judges' ratings was $r = .98$. On the basis of this information it was decided that Judge A's ratings of the remainder of the cases would stand in all instances in which she was confident. For those initial ratings of which Judge A was not confident, the correlation between the two judges' ratings was $r = .84$. On the basis of this information it was decided that conference ratings would be arrived at by two judges for all of the cases for which Judge A was not confident of her rating. Ratings arrived at under this system constitute the data.

selected. These pairs were observed during the free period of the day in kindergarten. The observations were coded in terms of a number of dependence-independence systems, and the two groups were compared for the frequency with which they exhibited behavior in each of these systems.

Subjects

For inclusion in this comparative study, any child had to meet the following criteria: (a) had been enrolled in his present public school kindergarten for at least three months; (b) had no siblings over 14 years of age; (c) was not a member of any minority group; (d) was born in the United States; (e) was from an intact (united) family. From the group of children who met these initial criteria, the WM Ss (working mothers' children) and the non-WM Ss (non-working mothers' children) were selected.

A child was included among the WM Ss if he met the following additional criteria: (a) mother was presently employed full-time, i.e., at least 32 hours a week; (b) mother had been so employed for at least the past six consecutive months. A child was included among the non-WM Ss if he met the following additional criterion: (a) mother had never worked during the lifetime of the subject. (A single work episode of less than a month's duration was not considered to be disqualifying.)

From the pool of potential subjects thus defined, matched pairs, each consisting of a WM S and a matched non-WM S, were selected for inclusion in the observation study according to the following matching criteria: (a) both had to be of same sex; (b) both had to be enrolled in same class at school; (c) the two had to be no more than four months different in age; (d) the two had to be no more than one level different in terms of social status of father's occupation; (e) the two had to belong to families of similar sizes, in terms of number of siblings; and (f) the two had to occupy similar ordinal positions within their own families.

The purpose of applying so many stringent criteria in the selection and matching of Ss was to obtain for comparison a group of WM Ss and a group of non-WM Ss who differed clearly in terms of working statuses and working histories of their mothers but who were directly comparable in terms of other variables which might be relevant to dependence and independence—variables such as age, sex, position in family, place of birth, intactness of family, social level of family, and minority group membership, as well as membership in the same class in school. In fact, the selection and matching criteria listed above represent a relaxation of our initial ones, which called for pairing of subjects holding identical ordinal positions and belonging to families of identical social level. Even with our somewhat relaxed criteria a great many children were disqualified from study, and many more could not be used because no suitable match was available.

In all, 26 matched pairs were obtained, 10 pairs of girls and 16 pairs of boys. In age the 52 Ss ranged from 63 months to 74 months, with an average age of 68.2 for the WM Ss and 67.6 for the non-WM Ss. In terms

of social status of father's occupation the range was from level 1 to level 6, with the WM Ss and the non-WM Ss each having a median level of 4. For number of siblings the range was from 0 to 5, with the WM Ss having 1.3 siblings on the average and the non-WM Ss having 1.7.³

Observation Method

Locale and time. All observations of Ss were made in the kindergarten classroom during the free activity period indoors which typically occurred early in the half-day school program. This period lasted at least 30 minutes. Each S was observed for four 15-minute periods, two of these during the first half of the activity period and two during the second half. A WM S and his matched non-WM S were observed on the same day, each for 15 minutes. The purpose of observing both members of a pair on the same day and of systematically alternating the order in which the two were observed was to render as equivalent as possible the situations in which the two members of the pair were observed and thus to clarify the meaning of any possible differences in the behavior of the two.

To assure variability in the observation situations, the four observation periods for any pair were spread over a period of several weeks, and each was set on a different day of the week. It was impossible to maintain constant intervals of time between observations of the various pairs, both because of occasional absences of Ss from school and because of the occurrence of special school events which rendered observation infeasible on certain days. Elapsed time between first and fourth observation for the various pairs ranged from 18 days to 58 days, with a mean of 32 days.

Recording of behavior. Detailed written accounts were used to record the child's behavior and pertinent information concerning the behavior of others as they interacted with him. The observer's record was divided into two columns, one for use in recording the S's behavior and the other for use in describing the situation, a technique which enables the observer to check constantly on the adequacy of her record. The observer attempted to obtain as comprehensive a record as possible, guided by knowledge of the coding system to be used and also by a guide for observers prepared by Stolz (9, pp. 342-343). The observer avoided interaction with the S or with other children and attempted to be unobtrusive in her behavior in the classroom.

Two observers participated in the study, one observing girls and the other observing boys. At no time did the two observers work simultaneously in the same classroom.

³ These differences in average number of siblings, although not large, do reflect the fact that working mothers typically have fewer children than do nonworking mothers—as was indicated in our survey data—and it is therefore very difficult to find exact matches in terms of number of siblings when other matching variables also must be considered. Nonetheless, all but two of the matched pairs of subjects either had the same number of siblings or differed by but one; the two exceptions each differed by two siblings.

CHILD DEVELOPMENT

Preparation of observation records. In preparation for the coding, the observation records were divided into units. The definition of a unit was adapted from Stevenson (8):

A behavior unit is that portion of behavior which contains: (a) an event (external or internal activity) which is psychologically different from events preceding and following; and (b) the subject's behavior in relation to the event, which is psychologically different from preceding and subsequent behavior (p. 32).

Two modifications were made of Stevenson's definition. First, goal-directedness of the behavior was emphasized, and second, provision was made for interrupted sequences of behavior—sequences in which a child's behavior toward a goal was interrupted and then resumed. Using the notion of overlapping episodes of Barker and Wright (2, p. 260), we provided for the inclusion of units whose segments were not continuous.

Coding. Each behavior unit, once identified, was classified into one of 10 categories. Nine of these categories represented different behavior systems relevant to dependence and independence; the tenth was a miscellaneous or residual category. The nine behavior systems were: *aggression, conformity, dominance, nurturance, obedience, self-reliance, sociability, submissiveness, succorance*. The basis for breaking down the dependence-independence system into less general categories is presented in Whiting *et al.* (11), to whom we are indebted. The present breakdown represents a revision of theirs, based on later work by B. B. Whiting and others⁴ as well as on our experiences in the course of the present study.

The relation of the eight systems to dependence and independence is, briefly, that the first system (succorance) is the sort of dependence most characteristic of infancy (but capable of occurring at all later ages as well), and that the other systems represent forms of socially relevant behavior which replace or supplement succorance at later ages and which we feel constitute a more useful classification than the simple dichotomy of dependence versus independence (11, p. 9).

For a discussion of these systems the reader may turn to (11); the details of their use in the present study are given in Hitchcock (3).

Each of the two observers divided her own records into units and then categorized these units. In addition, each determined whether or not each unit involved interaction with the teacher or another adult. This work was not initiated until all observation records had been collected, a procedure which eliminated the hazard that early experiences in encoding might alter subsequent practices in recording observations.

Scoring. The information contained in each child's four observation records was combined in the scoring. We determined the percentage of each child's units falling in each of the 10 categories and also the percentage

⁴ We are grateful to J. W. M. and B. B. Whiting for making their recent work available to us.

involving interaction with an adult. The percentages constituted the child's scores used in the statistical analysis.

Precautions against bias. The structure of the research plan incorporated certain procedures designed to eliminate the possibility that the data might be biased by anyone's beliefs concerning maternal employment and its effects on children. These precautions were:

1. The observer did not know which member of any pair was the WM S and which was the non-WM S. The selection and pairing of Ss was done by another worker than the observer, and the observer was told nothing of the Ss' backgrounds. Classroom teachers were instructed to avoid discussing the children's backgrounds with the observer, and the observers were prepared to remind the teachers to observe this precaution if a reminder were necessary. The observers remained "blind" with respect to their subjects until their analyses of their records were complete, i.e., until every child's scores were ready for statistical analysis.

2. The classroom teacher was not told which S was being observed at any time, nor was she told the nature of the behaviors of interest in the research. These two precautions with respect to the teacher, together with instructions to her to conduct her classes as she might in the absence of any observer, mitigated against the possibility that a teacher would treat any S in an unusual or special manner during the observation periods.

Reliability. The two observer-coders worked together closely in planning the details of the observation and coding procedures, a collaboration which undoubtedly contributed to the reliability of their independent work.

After each had had some practice in recording observations separately, the two made simultaneous but independent observations of the behavior of two boys, observed successively, in a school activity period. Later, after each had undergone detailed training and practice in the encoding procedure, they independently encoded and scored these boys' records. For the reliability analysis all of the scores obtained for one boy by one observer were ranked, and these ranks were correlated with the ranks of the scores obtained by the other observer for that boy. For one of the boys, this correlation was $\rho = .80$; for the other, it was $\rho = .74$ (7, pp. 202-213). These correlations represent interobserver agreement based on a single observation session. It is reasonable to think that such agreement would be higher for the data used, which are based on four observation sessions rather than one.

Due to various restrictions within which this study was conducted, including limitations of observer time, it was not possible for us to make a more thorough investigation of the reliability of our measures.

Results

Normative findings. For descriptive purposes we determined the average percentage of units which fell in each of the categories. These averages are shown in Table 1, given separately for boys and girls. Table 1 also shows

CHILD DEVELOPMENT

TABLE I
NORMATIVE FINDINGS: RELATIVE FREQUENCY OF OCCURRENCE OF
BEHAVIOR UNITS IN THE VARIOUS SYSTEMS

System of Behavior	Boys (N=32)		Girls (N=20)	
	Mean Per Cent of Units in System	Rank of System	Mean Per Cent of Units in System	Rank of System
Aggression	10.30	3	3.08	7
Conformity	6.71	6	5.97	4
Dominance	9.48	4	3.57	6
Nurturance	2.73	9	2.65	8
Obedience	9.13	5	11.36	2
Self-Reliance	4.96	8	5.81	5
Sociability	15.02	1	10.82	3
Submissiveness	5.40	7	2.06	9
Succorance	14.40	2	12.80	1
Miscellaneous (unclassifiable) ..	21.87	(1)	41.92	(1)

the rank of each average. Although the clear majority of all units were classifiable in one of the nine behavior systems, the tenth (miscellaneous) category contained more units than any single system. This was partly because many concomitant and interrupting behaviors could not be interpreted as to goal. As reference to Table 1 will reveal that, among boys, *sociability*, *succorance*, and *aggression* were the most frequently observed of the dependence-independence systems. For girls, *succorance*, *obedience*, and *sociability* were the most frequently observed.

Group differences. The principal analysis of the data was undertaken to discover differences between the children of working and nonworking mothers. Statistical tests were performed separately on the data for boys and the data for girls, as well as on the data for the sexes combined. For each system of behavior a difference score was obtained for each matched pair, and a *t* test on the mean difference was performed. The results for the sexes separately are shown in Table 2. Two-sided tests were used in each instance, and the probability associated with each value of *t* is shown if it is equal to or less than .10.

As reference to Table 2 will reveal, of the 20 tests which were performed, only one revealed differences at a level of significance widely acceptable to research workers. Under the assumption that the groups do not differ at all (i.e., under the null hypothesis), we expect to obtain one "significant" result for every 20 tests performed when we work at the .05 level. This chance expectation is confirmed in the present instance by the finding concerning girls' *interaction with adults*. We must conclude that, when the sexes are considered separately, no differences emerge between the children of working mothers and the children of nonworking mothers.

TABLE 2

GROUP DIFFERENCES BETWEEN THE CHILDREN OF WORKING AND NON-WORKING MOTHERS WITH THE SEXES CONSIDERED SEPARATELY

System	BOYS: 16 MATCHED PAIRS			GIRLS: 10 MATCHED PAIRS		
	<i>t</i>	Group Higher	<i>p</i>	<i>t</i>	Group Higher	<i>p</i>
Aggression	0.89	non-WM	..	1.09	WM	..
Conformity	0.54	non-WM	..	0.09	WM	..
Dominance	1.54	WM	..	1.58	WM	..
Nurturance	1.19	non-WM	..	0.76	WM	..
Obedience	1.04	WM	..	1.88	non-WM	<.10
Self-Reliance	1.03	non-WM	..	1.15	WM	..
Sociability	2.10	non-WM	<.10	0.08	WM	..
Submissiveness	0.43	non-WM	..	1.02	non-WM	..
Succorance	1.53	WM	..	1.70	non-WM	<.10
Interaction with Adults	1.52	WM	..	2.97	non-WM	<.02

On the possibility that the above over-all finding was due to the small sizes of the two samples, the data for the two sexes were combined, and *t* tests on the mean differences were performed for the 26 matched pairs thus obtained. The results are presented in Table 3. Again, we have no clear indication of group differences; it would be hazardous to attribute one "significant" result out of 10 (the apparent group difference in *dominance*—see Table 3) to any but chance factors.

TABLE 3

GROUP DIFFERENCES BETWEEN THE CHILDREN OF WORKING AND NON-WORKING MOTHERS WITH DATA FROM THE SEXES CONSIDERED TOGETHER

System of Behavior	<i>t</i>	Group Higher	<i>p</i>
Aggression	0.10	WM	..
Conformity	0.35	non-WM	..
Dominance	2.23	WM	<.04
Nurturance	0.10	non-WM	..
Obedience	0.45	non-WM	..
Self-Reliance	0.13	non-WM	..
Sociability	1.48	non-WM	..
Submissiveness	0.94	non-WM	..
Succorance	0.83	WM	..
Interaction with Adults	0.34	non-WM	..

CHILD DEVELOPMENT

A *post hoc* analysis of the data in Table 2 yielded a finding meriting attention in future research. Upon inspecting Table 2, the reader will notice that the group differences (between WM and non-WM Ss) are sometimes in the same direction for both boys and girls and are sometimes in the opposite direction. In *dominance*, for example, the WM boys score higher than the non-WM boys and also the WM girls score higher than the non-WM girls. Similarly, in *submissiveness* both the WM boys and the WM girls score lower than their like-sexed counterparts among non-WM Ss. On the other hand, in *aggression* the WM boys score lower than their matched pairs whereas the WM girls score higher than their matched pairs, an example of the group differences being in the opposite direction for the two sexes. Under the null hypothesis that only chance factors determine the slight group differences represented in Table 2, we would expect the two patterns to occur equally often. That is, we would expect the group differences to be in the same direction for the two sexes about as often as the group differences are in the opposite direction for the two sexes. However, only twice (for *dominance* and for *submissiveness*) were the group differences in the same direction for both sexes; for the other eight variables the group differences were in the opposite direction for the two sexes. This *post hoc* observation suggests that the implications of maternal employment for children, insofar as there are any implications at all, are different for boys than they are for girls. If the data of this study were amenable to a sophisticated statistical analysis, this would appear as an interaction between sex of child and employment status of mother. Whereas the differences between WM Ss and non-WM Ss are not themselves sufficiently great to be significant, the relative frequency with which these small differences are in the opposite direction for the two sexes is noteworthy.

It should be mentioned that none of the analyses we have presented has involved any direct comparisons of boys with girls. Such comparisons were precluded by the fact that the observations of girls were collected and encoded by a different person than the observations of boys so that any apparent sex differences would represent both sex and observer differences. Our analysis has been concerned entirely with difference scores for matched subjects of the same sex, an approach which eliminated any possible effects of observer differences on our findings.

SUMMARY AND CONCLUSIONS

In previous research evidence concerning differences between the children of working and nonworking mothers, differences associated with mother's employment status have been confounded by differences associated with other factors (family size, family intactness, family income level, etc.) which are themselves associated with mother's employment status. Because of this, it was impossible to determine whether observed differences between

the children of working and nonworking mothers reflected the implications of maternal employment for children or whether they reflected the implications of such other factors as divorce, small family size, minority group status, etc. The results of other studies are confounded also by vagueness in the definitions of "working mothers" and by lack of consideration of the work histories of nonworking mothers serving as subjects. The principal purpose of the present study was to eliminate these confoundings by the use of a matching design and then to determine whether the children of working and nonworking mothers differ with respect to behaviors related to dependence and independence.

Our analysis of the gross demographic characteristics of families of working and nonworking mothers, as revealed in replies to the initial questionnaire, revealed that on the average the families of working mothers differed in size, in age composition, and in intactness from the families of nonworking mothers. (Interestingly, none of these factors differentiated between full-time and part-time working mothers.) These findings from the total survey sample reinforced our belief in the importance of using a matching design in a comparison of the children of working and nonworking mothers. Therefore, in the subsample studied the subjects were matched on each of these variables as well as on others. In addition, clear and rather conservative criteria were used in selecting "working" and "nonworking" mothers.

Using a matched pairs design, we have not found differences between the children of working and nonworking mothers with respect to behavior systems related to dependence and independence. Those "significant" differences which did emerge in the analysis are so few that they probably represent chance factors only. Thus, we cannot reject the null hypothesis that working mothers' and nonworking mothers' children are from the same population with respect to dependence and independence. Although nonrejection of the null hypothesis must lead to weaker conclusions than rejection of it, one may surely conclude from these data that maternal employment *per se* is not the overwhelmingly influential factor in children's lives that some have thought it to be.

Needless to say, this conclusion applies only to the age group represented by our subjects. One cannot say what the findings might be if such a study were conducted with younger or older children. Moreover, we have no information concerning the implications of maternal employment during infancy and the earliest years of childhood, since we included among the WM Ss any child whose mother was currently employed full-time and had been so employed for the immediately preceding six months, if the child met the other criteria for inclusion in the study.

To the extent that a mother's working may have implications for her children, *post hoc* inspection of the data of this study suggests that the implications may be different for the two sexes. This finding may be interpretable in terms of the growing evidence that identification processes are different

CHILD DEVELOPMENT

in the two sexes (5), and it suggests the need for further research concerning maternal employment and sex-typing in children.

REFERENCES

1. ADAMSON, JEAN. A report of a study of kindergarten girls as part of an investigation into the effects of maternal employment. Unpublished master's paper in child development, Department of Education, Stanford Univer., 1958.
2. BARKER, R. G., & WRIGHT, H. F. *Midwest and its children*. Evanston, Ill.: Row Peterson, 1954.
3. HITCHCOCK, ETHEL A. The relationship between maternal employment and dependent behavior observed in kindergarten boys. Unpublished doctor's dissertation, Stanford Univer., 1958.
4. HOLLINGSHEAD, A. B. Index of social position. Mimeographed paper, Department of Sociology, Yale Univer., 1957.
5. LYNN, D. B. A note on sex differences in the development of masculine and feminine identification. *Psychol. Rev.*, 1959, 66, 126-135.
6. MACCOBY, ELEANOR E. Effects upon children of their mothers' outside employment. In *Work in the lives of married women*. (Proceedings of a Conference on Womanpower sponsored by the National Manpower Council.) New York: Columbia Univer. Press, 1958.
7. SIEGEL, S. *Nonparametric statistics for the behavioral sciences*. New York: McGraw-Hill, 1956.
8. STEVENSON, NANCY G. A method for analyzing observational records. Unpublished master's thesis, Stanford Univer., 1953.
9. STOLZ, LOIS M., *et al.* *Father relations of war-born children*. Stanford, Calif.: Stanford Univer. Press, 1954.
10. STOLZ, LOIS M., & ADAMSON, JEAN. Effects of mothers' employment on children: evidence from research. In preparation.
11. WHITING, J. W. M., CHILD, I. L., & LAMBERT, W., *et al.* *Field guide for the cross-cultural study of socialization*. New York: Social Science Research Council, 1953.
12. *Womanpower*. (A statement by the National Manpower Council). New York: Columbia Univer. Press, 1957.

ELECTROTACTUAL THRESHOLD IN THE NEONATE^{1,2}

LEWIS P. LIPSITT*

Brown University

and NISSIM LEVY

University of Portland

Several decades ago the Shermans (6, 7) reported on pain sensitivity in the human neonate, using a technique of pin-pricking areas of skin on different parts of the anatomy and noting the minimum necessary number of stimulations required to elicit withdrawal or avoidance responses. Their data indicated that pain sensitivity increases rapidly following birth and that parts of the body may be differentially sensitive, with the head area resulting in lower pain thresholds than the extremities. Pratt (5) has pointed out, however, that whether there is marked pain insensitivity in the newborn has been an unsettled question for some time, since several older studies failed to elicit pain reactions. Dockeray and Rice (1), for instance, found no change in responsiveness with increasing age; moreover, they reported an inverse cephalocaudal function for head, arm, and leg stimulation.

Most experimental procedures for determining pain sensitivity in the neonate until recently have utilized needle jabbing as the stimulating condition. Graham and her colleagues (2, 3) have used mild shock recently, observing the presence or absence of reaction to varying intensities of such stimulation in the knee region. Their data under these conditions have shown pain sensitivity as a function of age to follow the pattern found by the Shermans. The minimum necessary amount of stimulation required

* Walter S. Hunter Laboratory of Psychology, Brown University, Providence 12, Rhode Island.

¹ This investigation was supported by a research grant No. 3-B-9019 from the National Institute of Neurological Diseases and Blindness, Public Health Services, to the Institute for Research in Health Sciences at Brown University as part of a National Collaborative Project investigating neurological defects and their assessment in infancy. The authors are indebted to Dr. Glidden L. Brooks, director of the project at Brown, for facilitating this research in numerous ways. Thanks are also due the staff of the Providence Lying-In Hospital for their cooperation in this research.

² Stephen Klapper and Guenter Rose assisted in the collection of data.

CHILD DEVELOPMENT

(in volts) to elicit a withdrawal response of the stimulated leg decreases with increasing age during the first five days of life. Graham *et al.* (4) have further reported pain threshold comparisons between presumably normal infants and a group of stressed (i.e., condition of anoxia present at birth) infants, indicating significantly higher thresholds for the stressed than the normal subjects. It was concluded that this measure provides a possibly useful diagnostic procedure for the assessment of presence and severity of neurological damage in neonates.

The two studies reported here investigated electrocutaneous thresholds in neonates from birth through the fourth day of age, separately for males and females, under conditions both of successive daily testing of the same children and of single testing of each child at one day level only. The conditions of stimulation used and the method of threshold determination are somewhat different from those of Graham *et al.*, and one purpose of these studies was to provide normative data for the present stimulation procedures to be used later in the evaluation of effects of birth stresses.

METHOD

A Grass S-4 Physiological Stimulator with a stimulus isolation unit was used to provide the electric shock stimulus. Each electrode consisted of silver wire suspended in plastic tubing open at one end only. This tubing was filled with .9 per cent saline solution; the open tube end was easily pressed against the surface of the infant's skin and secured with adhesive tape from a flat plastic lip attached to the tubing. Both the indifferent and stimulating electrodes were kept in place throughout a day's test period, which lasted 15 to 20 minutes. Since the electrode wire itself did not touch the skin, and the stimulation was mediated by saline solution, constant surface contact was practically assured. The indifferent electrode was $\frac{1}{2}$ in. in diameter and placed on the back calf of the left leg, and the stimulating electrode was $\frac{1}{8}$ in. in diameter and attached to the large toe of that leg.

In all stimulations, a 10 c.p.s. current with monophasic pulse duration of 2 ms. was used, and voltage was varied from a low subthreshold value upwards to the intensity at which a reaction was elicited. Threshold for any given day's test period was the mean of three successive determinations or trials on that day, each such determination being made by administering an ascending series of shocks. On the first determination for any day, shock was administered first at 5 volts; then steps of 5 volts were used to the point of reaction.

The response was a rather discrete extension of the large toe (or all toes) often accompanied by withdrawal of the entire leg. Each stimulation was manually applied for a two-second period from the stimulator controls. Each of the three determinations on any day involved eliciting two successive responses at adjacent steps of 5 volts. That is, after the first reaction was noted, the stimulus was increased 5 volts and, if a second response

was obtained, the lower of the two stimulus values in volts was taken as the threshold value. If a second response was not obtained, intensity was increased until two responses to adjacent voltage settings was obtained. Under these conditions, close to perfect interobserver agreement may be achieved for the threshold intensity value. In the second and third such trials on a given day, the stimulus series was started 20 volts below threshold obtained on the previous trial.

STUDY I

The first study involved the testing of 36 neonates, 18 males and 18 females, longitudinally. Each infant was first tested within 24 hours of birth, then was given three additional tests at 24-hour intervals through the fourth day of life. Figure 1 presents the results of this study. A general decline in sensitivity threshold over the first four days may be seen for all Ss combined. In addition, a rather striking difference appeared between males and fe-

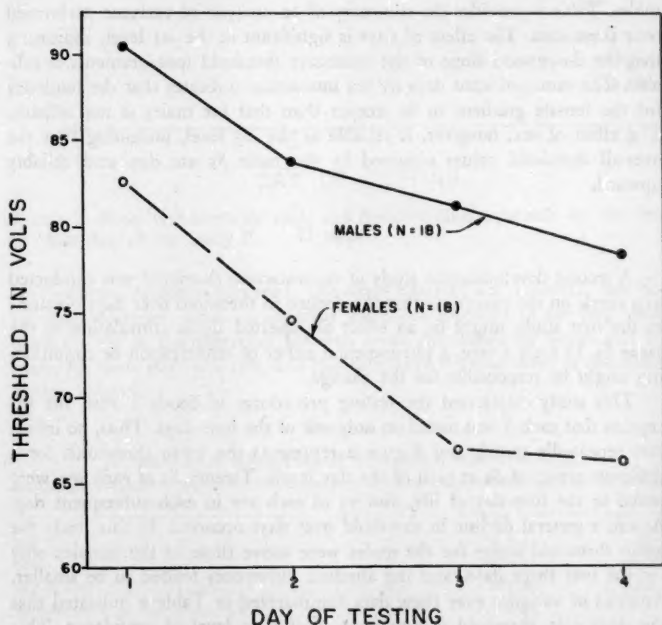


FIGURE 1—Mean thresholds for male and female infants separately for the first four days of life. Study I.

CHILD DEVELOPMENT

TABLE I

SUMMARY OF ANALYSIS OF VARIANCE FOR STUDY I

Source	df	ms	F
Between Ss	35		
Sex	1	4203.00	5.12*
Error (b)	34	821.44	
Within Ss	108		
Days	3	1449.00	12.89**
Days \times Sex	3	71.33	.63
Error (w)	102	112.44	
Total	143		

* $p < .05$.** $p < .01$.

males. Table I provides the summary of an analysis of variance performed over these data. The effect of days is significant at the .01 level, indicating that the downward slope of the successive threshold measurements is reliable. The nonsignificant days by sex interaction indicates that the tendency for the female gradient to be steeper than that for males is not reliable. The effect of sex, however, is reliable at the .05 level, indicating that the over-all threshold values obtained by the male Ss are displaced reliably upward.

STUDY II

A second developmental study of electrotactual threshold was conducted as a check on the possibility that the decline in threshold over days obtained in the first study might be an effect of repeated shock stimulation to the same Ss. In such a case, a phenomenon either of sensitization or conditioning might be responsible for the change.

This study duplicated the testing procedures of Study I with the exception that each S was tested on only one of the four days. Thus, no infant was repeatedly tested, and Figure 2 represents the mean thresholds for a different group of Ss at each of the day levels. Twenty Ss of each sex were tested in the first day of life, and 15 of each sex in each subsequent day. Again, a general decline in threshold over days occurred. In this study the mean threshold scores for the males were above those of the females only for the first three days, and the absolute differences tended to be smaller. Analysis of variance over these data summarized in Table 2 indicated that the decline in threshold is significant at the .01 level of confidence. This finding thus indicated that the threshold change occurring with age cannot be attributed to repeated testing of the same children. Again, no significant

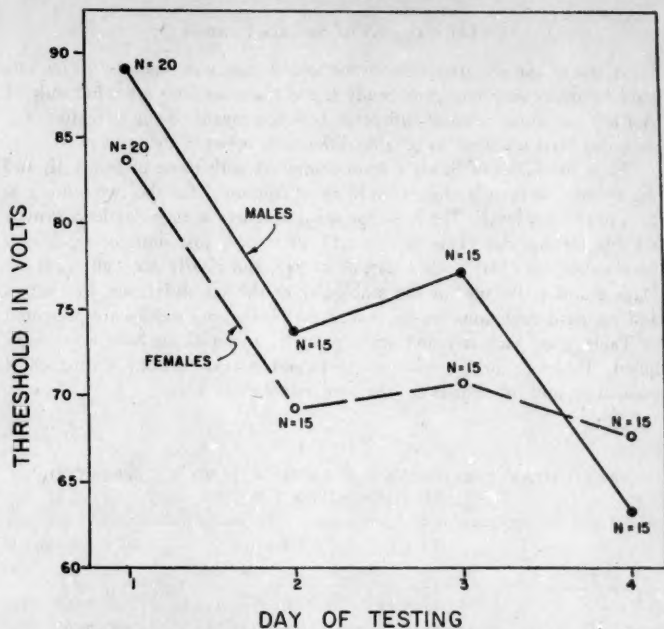


FIGURE 2—Mean thresholds for male and female infants separately for the first four days of life. Study II.

sex by days interaction occurred. The main effect of sex, however, was not reliable. Moreover, since an apparent sex effect appeared in the first three days, a t test of the difference between mean thresholds for males and females for these days was run, and this difference fell just short of reliability at the .05 level.

TABLE 2
SUMMARY OF ANALYSIS OF VARIANCE FOR STUDY II

Source	df	ms	F
Sex	1	326.43	1.07
Days	3	2887.70	9.48**
Sex \times Days	3	187.79	.61
Within cells	122	304.48	
Total	129		

** $p < .01$.

CHILD DEVELOPMENT

COMPARISONS OF STUDIES I AND II

A test of the comparability of the two studies was made by combining data for males and females in Study I, and those for both sexes in Study II. A *t* test for significance of difference between means of the two studies at each day level revealed no reliable differences between the studies.

Next the males of Study I were combined with those of Study II, and the females of both studies were likewise combined for the two studies at each of the day levels. The *t* test for sex differences at each day level proved reliable for one day (Day 3, $t = 2.28$, $df = 64$), just short of significant for another day (Day 1, $t = 1.97$, $df = 74$), and clearly not significant for Days 2 and 4. Because of the ambiguity of the sex difference, the means and standard deviations for the pooled data from both studies are presented in Table 3 for each day and sex separately, and also for both sexes combined. Table 3, then, constitutes threshold norms for our normal-child population and procedures for the first four days of life.

TABLE 3
POOLED MEAN THRESHOLDS FOR EACH AGE AND SEX SEPARATELY
AND FOR COMBINED SEXES

Age	MALE			FEMALE			COMBINED		
	N	M	SD	N	M	SD	N	M	SD
First day	38	89.95	13.9	38	83.24	15.2	76	86.59	15.0
Second day	33	79.12	16.2	33	72.15	17.9	66	75.63	17.4
Third day	33	79.30	17.3	33	68.70	19.8	66	74.00	19.4
Fourth day	33	71.42	17.9	33	66.91	18.0	66	69.16	18.1

DISCUSSION

These studies have demonstrated reliably that electrocutaneous sensitivity in the neonate increases sharply within the first four days of life. The data of both studies taken together indicate that this change in sensitivity is not attributable to the experimental conditions used to assess it. It is impossible to determine the physiological mechanism underlying this increasing sensitivity with age on the basis of the present study. One possibility is that thresholds are higher when the baby is asleep, the child tends to be awake an increasing amount of time on successive days, and the baby has a higher probability of being tested during an awake state with increasing age. Our running observations, while they suggest that thresholds are higher when the infants appear to be asleep, do not suggest that the infants tended to be awake more on successive testing days than earlier ones. Another possibility is that the infants are undergoing a slow recovery from effects of maternal anesthesia during the four hospital days. Still another possibility,

in line with the Graham (4) findings of higher thresholds for infants subjected to anoxia at birth, is that most or all infants undergo some degree of anoxia during the early hours or days of life due to the immaturity of the respiratory system and that all infants require time for the effects to dissipate. That infants recover at different rates from these respiratory traumas and that recovery rate is related to amount of residual damage or evidence of later impairment offer an intriguing possibility for further study. It is apparent that many parameters possibly relevant to sensitivity thresholds in the young child require systematic investigation.

It has been pointed out that divergent results have been obtained by various researchers dealing with pain sensitivity measures in the newborn. Pratt (5) has suggested that there has been much variation in the usage of the term pain. The view is taken here that pain is a construct whose presence is assessed on the basis of specific stimulating conditions and by inference according to arbitrary operational criteria of responsiveness. It is obvious from the Sherman, the Graham, and our data that whether pain responses occur in the neonate depends upon the intensity of stimulation administered. Furthermore, it seems clearer, although more cumbersome, to speak of our measure as a threshold of reaction to an aversive (electrotactual) stimulus; the nature of the response in the present studies does not suggest pain in the usual behavioral and neurological sense. The fact is that while all of our infants showed readily identifiable and discrete withdrawal responses of the stimulated leg, few showed general bodily discomfort at around threshold values of stimulation. If we had taken crying, grimacing, or some other responses as our pain criterion, on the other hand, the stimulating conditions reported here would not have been sufficient to reach threshold in most cases, and it is possible that the developmental function might have been altered somewhat.

SUMMARY

Two studies have been reported of the development of sensitivity to aversive shock stimulation in the child from birth through four days of age. Both studies revealed a reliable increase in sensitivity within this time period. One study produced a significant difference between the sexes, with the females reacting at lower intensity values of stimulation than the males. The second study did not reliably replicate this sex difference, although the means were in the same relationship to one another for the first three days of life. The combined data from both studies are presented as threshold norms for a population of normal infants under these conditions of stimulation.

REFERENCES

1. DOCKERAY, F. C., & RICE, C. Responses of newborn infants to pain stimulation. *Ohio State Univer. Stud. Contrib. Psychol.*, 1934, No. 12, 82-93.

CHILD DEVELOPMENT

2. GRAHAM, F. K. Behavioral differences between normal and traumatized newborns: I. The test procedures. *Psychol. Monogr.*, 1956, 70, No. 20.
3. GRAHAM, F. K., MATARAZZO, R. G., & CALDWELL, B. M. Behavioral differences between normal and traumatized newborns: II. Standardization, reliability, and validity. *Psychol. Monogr.*, 1956, 70 No. 21.
4. GRAHAM, F. K., PENNOYER, M. M., CALDWELL, B. M., GREENMAN, M., & HARTMANN, A. F. Relationship between clinical status and behavior test performances in a newborn group with histories suggesting anoxia. *J. Pediat.*, 1957, 50, 177.
5. PRATT, K. C. The neonate. In L. Carmichael (Ed.), *Manual of child psychology*. (2nd Ed.) New York: Wiley, 1954. Pp. 215-291.
6. SHERMAN, M., & SHERMAN, I. C. Sensorimotor responses in infants. *J. comp. Psychol.*, 1925, 5, 53-68.
7. SHERMAN, M., SHERMAN, I. C., & FLORY, C. D. Infant behavior. *Comp. Psychol. Monogr.*, 1936, 12, No. 4.

